

STIC-Biotech/ChemLib

86645

From: McElwain, Elizabeth
Sent: Wednesday, February 12, 2003 4:28 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

Please search for prior art and for interference
09/856,018 - SEQ ID NO: 15
and for DNA encoding SEQ ID NO: 16

Thank you,
Beth

Elizabeth F. McElwain, Ph.D.
U.S. Patent and Trademark Office
Tech Center 1600, Art Unit 1638
room CM1-9E11
mailbox CM1-9E12
703-308-1794
elizabeth.mcelwain@uspto.gov

CRFE

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/20
Date Completed: 2/24
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Result No.	Score			Query		DB	ID	Description
	Score	Match	Length					
1	1821	67.9	1942	4	US-09-326-203A-1			Sequence 1, Appli
2	802.5	29.9	1650	4	US-09-103-754A-2			Sequence 2, Appli
3	800.5	29.9	1376	3	US-09-165-042-2			Sequence 2, Appli
4	784.5	29.3	1766	4	US-09-326-203A-15			Sequence 15, Appl
5	784.5	29.3	1766	4	US-09-326-203A-16			Sequence 16, Appl
6	739.5	27.6	1895	4	US-09-326-203A-14			Sequence 14, Appl
7	568.5	21.2	629	4	US-09-103-754A-3			Sequence 3, Appli
8	396	14.8	4011	1	US-08-121-057-3			Sequence 3, Appli
9	396	14.8	4011	2	US-08-509-187D-3			Sequence 3, Appli
10	396	14.8	4011	2	US-09-121-396-3			Sequence 3, Appli
11	396	14.8	4011	5	PCF-US93-09704A-3			Sequence 3, Appli
12	390	14.6	2040	3	US-09-165-042-4			Sequence 4, Appli

Db 237 ATGGCGATTTGGATTCT---GCTGGCGTTACTACGGTGCAGGAAACGGTGGCGGAGAG 293
Qy 19 -----LeuArgArgProSer-----AlaThrSerThrAlaGlyLeu 31
Db 294 TTGCTGCGATCTTGATAGCTTCGTCGACGGAAATCGAGATCGGATCTTCTCAACGGACTT 353
Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspLeuAlaLysSerGly 51
Db 354 CTCTCTCTGCTCCGATAAATCTCTCTCGGATGATGTGGAGCTCCCGCGGACGTT 413
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67
Db 414 AGGATCGGATGATTCGTTTAAACGATCAGCTCAGGGAACAGCCATTTGGCCGGA 473
Qy 68 GlnAsnGlu-----LysGln 72
Db 474 GATAATAACGGTGTGGCGATAATAACGGTGTGGAGAGCGCGGAGAGAGGA 533
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
Db 534 AACCCGATGCTAGC-----TTTACGTTATCGACCGCTCCAGCTTCATCGGAGG 584
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112
Db 585 CGGAGAGAGAGTCCATGCTCGACGCAATCTTCAACAGAGCCATGCCGGAATATTC 644
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132
Db 645 AACCTCTGTGTAGTAGTTCTTATTGCTGTAACAGTAGTACATCATCGAATACTTATG 704
Qy 133 LysTyrGlyTrpLeuLysSerGlyPheThrPheSerSerLysSerLeuArgAspTrp 152
Db 705 AAGTATGGTGTGATCAGAACGAGATTTCTGGTTTATGTTCAAGATCGCTCGGAGATTGG 764
Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172
Db 765 CCGCTTTTCATGCTGTGTATATCCTTTCGATCTTCTTGGCTGCCCTTACCGTTGAG 824
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValLeuHisIleIleIle 192
Db 825 AAATTGGTACTTCAGAAATACATATCAGAACCTGTGTGTCATCTTCTTCATATTATTATC 884
Qy 193 ThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
Db 885 ACCATGACAGAGGTTTGTATCCAGTTTACCTCACCTAACGCTGATTCGCTTTTATA 944
Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232
Db 945 TCAGGTGTCACTTTGATGCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1004
Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
Db 1005 CATACTAGCTATGACATAAGATCCCTAGCCCAATGCAGCTGATAGGCC----- 1052
Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272
Db 1053 -----AATCCTGAAGTCTCCTACTAGCTTAGCTTGAAGAGCTTGGCATATTTTCATG 1103
Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292
Db 1104 GTCCCTCCACATGTTGTTATCAGCCCAAGTTATCCAGCTTCTGCTATACGGAAGGTT 1163
Qy 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312
Db 1164 TGGTGGCTGCTCAATTTGCAAAACTGTCATATTCACCGGATTCGGGATTTATAATA 1223
Qy 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
Db 1224 GAACAATATAAATCTTATGTCAGGAACCTCAAGATCCTTTGAAAGCGGATCTCTTA 1283
Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
Db 1284 TATCCTATTGAAACAGTGTCAAGCTTTCAAGCTTTCAGTTCCAAATTTATATGTGCTCTGCATG 1343

Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
Db 1344 TTCCTACTCTCTTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTCTGCTTCGGGAT 1403
Qy 373 ArgGluPheTyrLysAspTrpPheAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
Db 1404 CGTGAATCTTCAAAAGATTGGTGAATGCAAAAGTGTGGAGATTACTGGAGAAATGTGG 1463
Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
Db 1464 AATATGCTGTTCATAAATGGATGTTCCGACATATATATCTCCCTGCTTCCGACGCAAG 1523
Qy 413 LeuProLysAlaAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1524 ATACCAAGACACATCGCCATTATCATCTTCTTAGTCTCTGCAGTCTTTCATGAGCTA 1583
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
Db 1584 TGCATCGCAGTTCCTTGTCTCTCAAGCTATGGGCTTTTCTTGGGATTTATGTTTCAG 1643
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1644 GTGCTTTGGTCTTCATCACAACATCTCTCAGGAAAGGTTT---GGCTCAACGGTGGGG 1700
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
Db 1701 AACATGATCTTCTGGTTCATCTTCTGCATCTTCTGCATTTTCGACAAACGATGTGTGCTCTCTTAT 1760
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1761 TACCACGACCTGATGAACCGAAAGGATCGATG 1793

RESULT 2

US-09-103-754A-2

; Sequence 2, Application US/09103754A

; Patent No. 6344548

; GENERAL INFORMATION:

; APPLICANT: Farese, Robert

; APPLICANT: Cases, Sylvaine

; APPLICANT: Smith, Steven

; APPLICANT: Erickson, Sandra

; TITLE OF INVENTION: diacylglycerol O-acyltran

; TITLE OF INVENTION: sferase

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/103,754A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Field, Bret E

; REGISTRATION NUMBER: 37,620

; REFERENCE/DOCKET NUMBER: 6510-105p

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650 327 3400

; TELEFAX: 650 327 3231

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1650 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-103-754A-2

Alignment Scores:

Pred. No.: 1.64e-82 Length: 1650
 Score: 802.50 Matches: 189
 Percent Similarity: 55.16% Conservativity: 89
 Best Local Similarity: 37.50% Mismatches: 175
 Query Match: 29.94% Indels: 51
 DB: 4 Gaps: 16

US-09-856-018B-16 (1-504) x US-09-103-754A-2 (1-1650)

```

QY 17 SerSerLeuArgArgPro---SerAlaThrSerThrAlaGlyLeuPheAsnSerPro 35
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 ASCTCTGGCGTGGAGGACCGGCTCGCGGGTTTCGTCCAGGGT----- 164

QY 36 GluThrThrAspSerSerGlyAspLeuAlaLysAspSerGlySerAspSer 55
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 165 -----GGTAGTGGCCCAAGGTAGAAGAGGACGAGGTGCGAGACGGGCT 209

QY 56 IleAsnSerAsp-----AspAlaAlaValAsnSerGlnGlnAsnGlu 70
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 210 GTGAGCCCGGCTGGCGCCGGGGTGACGCGCGGCTCCGGCTCCAGCCCAT 269

QY 71 LysGlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerVal----- 87
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 270 ACCGGGACAAAGAC-----GGCGGACCAAGCGCTGGCGGACGCGC 308

QY 88 -----ProAlaHisArgLysValLysGluSerProLeuSerSerAspThrIle 103
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 TACTGGGATCTGAGGTGCCATCGT---CTGCAAGATCTTTGTTACGCTCAGACAGTGT 365

QY 104 PheArgGlnSerHisAlaGlyLeuPheAsnLeuLysIleValValLeuAlaValAsn 123
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 TTC---AGCAATTATCGTGGTATCTGGAATGGTGTGGTGTGATCGTCTGAGTAA 422

QY 124 SerArgLeuIleGluAsnLeuMetLysTyrGlyTrpLeuLysSerGlyPheTrp 143
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 GCAGGTTATTTTAGAAGACCTTATCAAGTATGGCATCTCTGGTG---GATCCTATCCAG 479

QY 144 PheSerLysSerLeuArgAsp-----TrpProLeuPheMetCysCysLeuSer 160
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GTGGTCTCTCTTTTGAAGGACCCCTACAGTGGCTGCCCTGCCCATGCGTGATTGCA 539

QY 161 LeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 TCAATATTTTGTGTGCTGCATTTCAAGTTGAGAGCGCCTGGCAGTGGTGCCCT 599

QY 181 ProGluProValValValLeuHisIleIleThrSerThrSerLeuPheTrp 200
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 600 ACAGAGCAGATGGGCTGCTGCTACATGGTTAACTGGCCACCAATCATTTGCTTCCA 659

QY 201 ValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPhe 220
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 660 GCAGCTGTGGCTTACTGGTTGAGTCTATCACTCCAGTGGGTTCCGTGTTGCTCTGGCA 719

QY 221 Ser---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyr----- 236
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 720 TCATACTCCATCATGTTCTCAAGCTTTATCTCCCGGGATGTCAACCTGTGGTGGCCG 779

QY 237 AspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeu--- 255
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 780 CAGCGAAGGCTCAAGCCCAAGCTGCTCTACAGGAGAGAAAGTCAAGTGGGCTGCTGCC 839

QY 256 -----AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVal 273
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 840 CAGCAAGCTGTGAGCTATCCACACAACCTGACCTACCGAGATCTCTATTACTTCATCT 899

```

Alignment Scores:

Pred. No.: 3.75e-82 Length: 1976
 Score: 800.50 Matches: 190
 Percent Similarity: 55.58% Conservativity: 89

RESULT 3

US-09-165-042-2

; Sequence 2, Application US/09165042

; Patent No. 6100077

; GENERAL INFORMATION:

; APPLICANT: Sturley, Stephen L.

; APPLICANT: Oelkers, Peter

; TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL

; FILE REFERENCE: ACYLTRANSFERASE

; CURRENT APPLICATION NUMBER: US/09/165,042

; CURRENT FILING DATE: 1998-10-01

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1976

; TYPE: DNA

; ORGANISM: Yeast

US-09-165-042-2

Db	60	AGCTCTCGGCGTCGGAGGACCGGCTCGCGGGTTTCCATCCAGGGTGGTAGTGGGCCCATG	119
Qy	32	PheAsnSerProGluThrThrAspSerSerGlyAspLeuAlaLysAspSerGly	51
Db	120	GTAGACGAGAGGAGTGGCAGACCGCTGTGGGCCCGGACTGGGC-----GCCGGG	173
Qy	52	SerAspAspSerIleAsnSerAspAlaLysValAsnSerGlnGlnAsnGluLys	71
Db	174	GGTGAGCGTCCGGCT-----CCGGCTCCGGTTCGGGTCCAGCCACACCCGGGAC	224
Qy	72	GlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArg	91
Db	225	AAAGACGGCAGACACCGTGGGCGACGCCACTGG-----GAGCTGAGTGGCCATCGT	278
Qy	92	LysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeu	111
Db	279	--CTGCAAGACTCTTTGTTCAAGCTCAGACACGGCTTC--AGCAATTACCGTGGTATC	332
Qy	112	PheAsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeu	131
Db	333	CTGAATTGGTGGTGTGATGTCCTGAGTAATGCAAGGTATTTTATAGAAATCTT	392
Qy	132	MetLysTyrGlyTrpLeuLysSerGlyPheTrpPheSerSerLysSerLeuArgAsp	151
Db	393	ATCAGTATGGCATCTCGTG--GATCCATCCAGGTGGTGTCTCTTTCTGAAGGAC	449
Qy	152	-----TrpProLeuPheMetCysLysLeuSerLeuValValPheProPheAlaAla	168
Db	450	CCCTACAGCTGGGCTCCCCCATCGTTGATCATTTGATGATCCATATCTTTATTGTGGCTACA	509
Qy	169	PheIleValGluLysLeuAlaGlnArgLysCysIleProGluProValValValLeu	188
Db	510	TTTCAGATTGAGAAGCGCCTGTCACTGGGTGCCCTGACAGACAGATGGGCTGCTGCTA	569
Qy	189	HisIleIleIleThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAsp	208
Db	570	CATGTGGTTAACTGGCCCAATATCTGTCTCCACAGCGTGGCCCTACTGGTGGAG	629
Qy	209	SerAlaPheValSerGlyValThrLeuMetLeuPheSer--CysValValTrpLeuLys	227
Db	630	TCTATCACTCCAGTGGGTTCCCTGTTTGTCTGGGATCATACTCCATCATCTCTCAAG	689
Qy	228	LeuValSerTyrAlaHisThrAsnTyr-----AspMetArgAlaLeuThrLysLeu	244
Db	690	CTTTTCTCTACCGGATGTCATCTGTGTGCCGCCAGCGGAAGGTCACAGGCCAAAGCT	749
Qy	245	ValGluLysGlyCluAlaLeuLeuAspThrLeu-----AsnMetAspTyrProTyr	261
Db	750	GTGTGCGAGGGAAGAGTCAAGTGGGCTGCTGCCAGAACACTGAAGTATCCCGGAC	809
Qy	262	AsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnPro	281
Db	810	AACCTGACCTACCGAGATCTCTATTACTTCACTTTTGCTCTACTTGTGTGTATGAATC	869
Qy	282	SerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeu	301
Db	870	AACTTCTTCGATCCGCCGAATACAAAGCGCTTCTGCTACGGCGGTTCTTGAGATG	929
Qy	302	IleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGln	321
Db	930	CTCTTTTTCACCCAGCTTCAGTGGGGGTGATCCAGCAGCTGGATGGTCTCTACTATCCAG	989
Qy	322	AspSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla-----ThrGluArgValLeu	339
Db	990	AACTCCATGAGCCCTTCAAG--GACATGGNACTATTCCAGCAATCATTTGACGGTCTCTTA	1046
Qy	340	LysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeu	359
Db	1047	AAGCTGGCGGTCCCCAACCATCTGATATGGCTCATCTTCTTCTATTTGGCTTTTCCACTCA	1106
Qy	360	TrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrp	379
Db	1107	TGTCCTAAATGCTGTGGCAGAGCTCTGCAAGTTTGTGGACACCGAGTCTCTACAGGACTGG	1166

```

Qy 380 TrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrp 399
Db 1167 TGGAAATGCTGAGTCTACCTTACTTTTGGCAGAACTGGAATATCCCGCTGCACAAGTGG 1226
Qy 400 MetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeu 419
Db 1227 TGCATCAGACACTTCTACAGGCTATGCTCAGACTGGGCGAGCAACAAATGGATGGCCAGG 1286
Qy 420 LeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHis 439
Db 1287 ACTGGGGTCTTTTGGCGTCACGCTTCTTCCATGAGTACTAGTAGCATTCCTCCCTGAGG 1346
Qy 440 IlePheLysLeuTrpAlaPheGlyIleMetPheGlnValProLeuValLeuIleThr 459
Db 1347 ATGTTTCCCGCTCTGGGCATTTACAGCCATGATGGCTCAGTCCAGTCCCACTGGCC 1397
Qy 460 AsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIle 479
Db 1398 ---TGGATGTGAACCGCTTCTTCCAGGGAAGTATGGCAATGCAGCTGTGTGG---GTG 1451
Qy 480 PheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAsp 495
Db 1452 ACACATCATCATTTGGGCAACCGGTGGCTGTGCTCATGTATGTCCACGAC 1499

RESULT 5
US-09-326-203A-16
; Sequence 16, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Lussner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleo
; TITLE OF INVENTION: Acid Sequences
; FILE OF INVENTION: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-326-203A-16

Alignment Scores:
Pred. No.: 2,18e-80 Length: 1766
Score: 784.50 Matches: 186
Percent Similarity: 55.85% Conservative: 91
Best Local Similarity: 37.50% Mismatches: 186
Query Match: 29.27% Indels: 33
DB: 4 Gaps: 16

US-09-856-018B-16 (1-504) x US-09-326-203A-16 (1-1766)
Qy 17 SerSerLeuArgArgArgPro---SerAlaThrSerThrAlaGly-----Leu 31
Db 60 AGCTCTCGGCGTCGAGGACCGGCTCGGGTTTCCATCCAGGTGGTAGTGGGCCCATG 119
Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysaspSerGly 51
Db 120 GTAGACGAAGAGGAGGTGCGAGACGCCGCTGTGGCCCGCATTTGGGC-----GCCGGG 173
Qy 52 SerAspAspSerIleAsnSerAspAlaAlaValAsnSerGlnGlnAsnGluLys 71
Db 174 GGTGACGCTCCGGCT-----CCGGCTCCGGTTCCGGCTCCAGCCACACCCGGGAC 224
Qy 72 GlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArg 91
Db 72 GlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArg 91

```

Db 225 AAAGACCGGCAGACCGCTGGCGGACGCCACTGG-----GAGCTGAGGTGCCACTCT 278
Qy 92 LysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeu 111
Db 279 ---CTGCAAGACTCTTGTTCAGCTCAGACACGGGTTTC---AGCAATTACCGTGGTATC 332
Qy 112 PheAsnLeuCysIleValValValLeuAlaValAsnSerArgLeuIleLeuGluAsnLeu 131
Db 333 CTGAATTGGTGGTGGTGTGATGCTGATCCTCGAGTAATGCAAGGTATTATTTAGAGAACTTT 392
Qy 132 MetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAsp 151
Db 393 ATCAAGTATGGCATCTCTGGT---GATCCCATCCAGGTGGTCTCTGTTTCTGAAGGAC 449
Qy 152 -----TrpProLeuPheMetCysLysLeuSerLeuValValPheProPheAlaLa 168
Db 450 CCTACAGCTGGCTGCCCTCCCTCATGTCATGTCATCAATATCTTTATTGGGTACA 509
Qy 169 PheIleValGluLysLeuAlaGlnArgLysCysIleProGluProValValValLeu 188
Db 510 TTTCAAGATTGAGAAGCGCTGTCACTGGTGGCTGCACAGACAGATGGGGCTGCTGCTA 569
Qy 189 HisIleIleIleThrSerThrSerLeuPheTrpProValLeuValIleLeuArgCysAsp 208
Db 570 CATGTGGTTAACTGGCCCAATATCTCTCTCCAGCAGCTGTGGCCCTACTGTTGAG 629
Qy 209 SerAlaPheValSerGlyValThrLeuMetLeuPheSer---CysValValTrpLeuLys 227
Db 630 TCTACATCCAGTGGGTTCCCTGTTGCTGTCGTCATCATCTCCATCATCTTCCCTCAG 689
Qy 228 LeuValSerTyrAlaHisThrAsnTyr-----AspMetArgAlaLeuThrLysLeu 244
Db 690 CTCTTCTCTACCGGATGTCATCTGTTGTCGCGCAGCGAGGTCAGGCCAAAGCT 749
Qy 245 ValGluLysGlyGluAlaLeuLeuAspThrLeu-----AsnMetAspTyrProTyr 261
Db 750 GTCTGTCAGGGAAGAAGTCACTAGTGGGCTGTCGCCCAAGCACTGAAGCTATCCGGAC 809
Qy 262 AsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnPro 281
Db 810 AACCTGACCTACCGAGATCTCTATTACTTATCTTCTGCTTCTTCTTGTGTTATGAACTC 869
Qy 282 SerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeu 301
Db 870 AACTTCTCTGATCCCGCAATACAAAGCGCTTCTGCTACGGCGGTTCTTCTGAGATG 929
Qy 302 IleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGln 321
Db 930 CTCTTTTTCACCCAGCTTCAAGTGGGCTGATCCAGCAGTGGATGCTCTACTATCCAG 989
Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla-----ThrGluArgValLeu 339
Db 990 AACTCCATGAAGCCCTTCAAG---GACATGGACTATTTCAGCAATCATTTGAGCGTCTCTTA 1046
Qy 340 LysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPheHisLeu 359
Db 1047 AAGCTGGCGGTCCTCAACCATCATGATATGGCTCATCTTCTTCTTATGGCTTTTCCACTCA 1106
Qy 360 TrpLeuAsnIleLeuAlaGluLeuArgPheGlyAspArgGluPheTyrLysAspTrp 379
Db 1107 TGTCTCAATGCTGGCAGAGCTCTGTCAGTTTGAGACCGCGAGTTCTACAGGACTGG 1166
Qy 380 TrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrp 399
Db 1167 TGAATGCTGAGTCTGTCACTTCTTTGGCAGAACTGGAATATCCCGTGGCACAAGTGG 1226
Qy 400 MetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeu 419
Db 1227 TGCATCAGACACTTCTACAAGCTTATGTCAGACTGGGCGAGCAACAATGGATGGCCAGG 1286
Qy 420 LeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHis 439
Db 1287 ACTGGGCTCTTTTGGCTCAGCCCTTCTTCCATGAGTACCTAGAGCATTCCTCCCTGAGG 1346

Qy 440 IlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThr 459
Db 1347 ATGTTCCGCTCTGGCATTCACAGCATGATGGCTCAGTCCACTGGCC-----1397
Qy 460 AsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIle 479
Db 1398 ---TGCATTGTGAACCGTCTCTTCCAAGGAACATATGCAATGCAGCTGTGTGG---GTG 1451
Qy 480 PheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAsp 495
Db 1452 ACACATCATTTGGGCAACGGTGGCTGCTCATGATATGTCACGAC 1499
RESULT 6
US-09-326-203A-14
; Sequence 14, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Lassner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1895
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (209)
; OTHER INFORMATION: n at position 209 is unknown
US-09-326-203A-14

Alignment Scores:
Pred. No.: 3,78e-75 Length: 1895
Score: 739.50 Matches: 158
Percent Similarity: 59.53% Conservative: 70
Best Local Similarity: 41.25% Mismatches: 138
Query Match: 27.59% Indels: 17
DB: Gaps: 9

US-09-856-018B-16 (1-504) x US-09-326-203A-14 (1-1895)

Qy 123 AsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPhe 142
Db 243 AATGCCGGTATTCTTCTGGAGAACCTCAAGTATGGCATCTCTGGTG---GACCCCATC 299
Qy 143 TrpPheSerSerLysSerLeuArgAsp-----TrpProLeuPheMetCysLysLeu 159
Db 300 CAGGTGGTTTCTCTGTCCTCTGAAGGATCCCTATAGTGGCCGCCCATGCTGGTTATT 359
Qy 160 SerLeuValValPheProPheAlaPheIleValGluLysLeuAlaGlnArgLysCys 179
Db 360 GCGGCCAATGTCTTTTCTGCTGCTGTCATTCAGTTTGAAGCGCTCGGCTGGTGCC 419
Qy 180 IleProGluProValValValValLeuHisIleIleIleThrSerThrSerLeuPheTyr 199
Db 420 CTGACGGACACGGGAGCTGCTGTCAGCTGGCCCACTGGCCCACTTCTGTGTTTC 479
Qy 200 ProValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeu 219
Db 480 CCAGCGGCTGTGTTCTTACTGTTGAGTCTATCATCTCCAGTGGGCTCCCTGCTGGCGCTG 539
Qy 220 PheSer---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsn---TyrAsp 237
Db 220 PheSer---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsn---TyrAsp 237

Db 540 ATGGCGCACACCACTCTCTCTCCTCAAGCTCTTCTCTACCCGAGCTCAACTCATGGTGC 599
Qy 238 MetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeu----- 255
Db 600 CGCAGGCGCAGGCGCAAGGCTGCTCTGCAGGAAGAAGCGCAGCTGCTGCTGCCCGC 659
Qy 256 ---AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAla 274
Db 660 CACACCGTGAGTACCGGCAATCTGACCTACCGGATCTCTACTCTCTCTCTCTCGCC 719
Qy 275 ProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeu 294
Db 720 CCCACCTTGCTGACGAGCTCAACTTCCCGCTCTCCCGCATCCGGAACGCTTCTTG 779
Qy 295 PheArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheIleIleAspGln 314
Db 780 CTGGCAGGATCTCTGAGTCTCTTCTACCCAGCTCCAGGTGGGCTGATCCAGCAG 839
Qy 315 TyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla 334
Db 840 TGGATGTGCCACACCACTCCAGACTCCATGAGCCCTTCAAG---GACATGGACTACTCA 896
Qy 335 -----ThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
Db 897 CGCATCATCGAGCGCTCTCTGAACTGGCGTCCCAATCACTCATCTGCTCATCTTC 956
Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
Db 957 TTCACCTGGCTCTCCACTCTCTGCTGAATGCCGTGAGCTCATGCAGCTTTGGAGAC 1016
Qy 373 ArgGluPheTyrLysAspTyrTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
Db 1017 CGGGAGTCTACCGGGGACTGTGGAATCCGAGTCTGTCACTCTCTCTGCGCAACTGG 1076
Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
Db 1077 AACATCCTGTGCACAGTGTGTGATCAGACACTTACAGCCCATCTCTCGAGGGGC 1136
Qy 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1137 AGCAGCAAGTGGATGGCGCAGGAGGGGTCTCTGCGCTCGGCTCTCTCCACGAGTAC 1196
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
Db 1197 CTGTGAGCGTCCCTCTGCGAATCTTCGCGCTCTGCGGCTTTCAGGGCATATGGCTCAG 1256
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1257 ATCCCACTGGCC-----TGGTCTGCGGCGCTTTTTCAGGGCACTATGGC 1304
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuTyr 492
Db 1305 AACGACGCTGTGG---CTGCTGCTCATATCGGACGCAATAGCGCTCTCTCATGTAC 1361
Qy 493 TyrHisasp 495
Db 1362 GTCCACGAC 1370

RESULT 7

US-09-103-754A-3

: Sequence 3, Application US/09103754A

: Patent No. 634548

: GENERAL INFORMATION:

: APPLICANT: Farese, Robert

: APPLICANT: Cases, Sylvaine

: APPLICANT: Smith, Steven

: APPLICANT: Erickson, Sandra

: TITLE OF INVENTION: Diacylglycerol O-acyltran

: TITLE OF INVENTION: sferase

: NUMBER OF SEQUENCES: 6

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Bozicevic & Reed

: STREET: 285 Hamilton Avenue, Suite 200

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/103,754A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 6510-105p
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650 327 3400
TELEFAX: 650 327 3231
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-103-754A-3

Alignment Scores:

Pred. No.: 3,2e-56 Length: 629

Score: 568,50 Matches: 105

Percent Similarity: 87,02% Conservative: 9

Best Local Similarity: 80,15% Mismatches: 13

Query Match: 21,21% Indels: 4

DB: 4 Gaps: 1

US-09-856-018B-16 (1-504) x US-09-103-754A-3 (1-629)

Qy 289 IleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIleLeuPheThrGlyValMet 308
Db 8 ATACGGAAGGTTGGTGGCTCGTCAATTGCAAAACTGGTCAATATTCACCGGATTCATG 67
Qy 309 GlyPheIleIleAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLys 328
Db 68 GGATTTATATAGAACATATATAAATCTATTGTCAGGAACCTCAAGCATCTCTTTGAAA 127
Qy 329 GlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrVal 348
Db 128 GGCGATCTCTATATGCTATTGAAAGAGTGTGAAGCTTTTCAGTTCCAAATTTATATGTG 187
Qy 349 TrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeu 368
Db 188 TGGCTCTGCTACTCTACTGCTTCTCCACCTTTGGTTAAACATATTCGACAGAGCTTCTC 247
Qy 369 ArgPheGlyAspArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyr 388
Db 248 TGGTTCGGGGATCGTGAATTTACAAAGATTGGTGGAAATGCTGGAAGATTGGGAGATTAC 307
Qy 389 Trp-ArgMetTrpAsnMetProValHisLysTrp-MetIleArgHisLeuTyrPhePro 407
Db 308 TGGGAGAATGTGGAATATGCTGCTCCATAATGGGATGGGTCGCGACATATATACCTTCCC 367
Qy 408 ---CysLeuArgHisGlyLeuProLys 415
Db 368 CGTGTTCGCGCACAGGATTACCCAAA 394

RESULT 8

US-08-121-057-3

Sequence 3, Application US/08121057
Patent No. 5484727
GENERAL INFORMATION:
APPLICANT: CHANG, TA-YUAN
APPLICANT: CHANG, CATHERINE C. Y.
TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL
TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,057
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: LAMPORT HAMMITTE, ANN.
REGISTRATION NUMBER: 34,858
REFERENCE/DOCKET NUMBER: DCI-033CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-2700
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4011 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1397..3046
US-08-121-057-3

Alignment Scores:
Pred. No.: 5,21e-35 Length: 4011
Score: 396.00 Matches: 150
Percent Similarity: 44.6% Conservative: 80
Best Local Similarity: 29.1% Mismatches: 191
Query Match: 14.7% Indels: 95
DB: 1 Gaps: 24

US-09-856-018b-16 (1-504) x US-08-121-057-3 (1-4011)

QY 33 AnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaLy 48
DB 1537 GATGCAAGAGATAAAGATTGACAGCAGCAGAGGAAATGAGCCATTTTATGAA 1596
QY 48 sAspSerGlySerAspSerIleasnSerAspAspAlaAlaValAsnSerGlnGlnG 68
DB 1597 GGAAGTTGGCAGT-----CACTTTGATGATTTTGTGACCAATCTCATTTGAAA 1644
QY 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79
DB 1645 GTCAGCATCATTAGATAAATGGTGGGCGCTCTCACAACTTTCTGTCTTGAAGAGA 1704
QY 80 -LysPheAlaTyrArgProSer-----ValPro----- 88
DB 1705 GAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1764
QY 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerH 108
DB 1765 TGAAGGCCCTCTCTTATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1810

QY 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeu 127
DB 1811 -AGACAATATATACATCTTTATGCCCTCTCTCATCTCTTTATCTCTCAGCACACATGT 1869
QY 127 eileGlu-AsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerL 147
DB 1870 AGTAGATTACATTGATGAAGAGGCTGTGCTTGAAGT-----TCAGCCT 1914
QY 147 ysSerLeuArg-AspTrpProLeuPheMetCysCysLeuSerLeuVal----- 162
DB 1915 CTTGCTTATGCTTTTGGCAAAATTTCTACCGTTGTTTGGACCTGGTGCATCATGTTCCT 1974
QY 163 ValPheProPheAlaAla---PheIleValGluLysLeuAla-----GlnArg 177
DB 1975 GTCTACATTTTTCAGTTCCCTATTTTCTGTTTCAACATTTGGCGCACCTGGCTATAGCAAG 2034
QY 178 LysCysIleProGluProValVal-----ValValLeuHisIleIleIleThrSer 194
DB 2035 TTCTCATCGCTGATCCGTTCTCTCTCTCATGGCTTTCTTTTCATGATCTTCCAGATTGG 2094
QY 195 ThrSerLeuPheTyr-ProValLeuValIleLeuArgCysAspSerAlaPheValSerG 214
DB 2095 AGTTCTAGCTTTTGGACCAACATATGTTGTTAGCATATACACTCCACACCTTCCCG 2154
QY 214 yValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHis 234
DB 2155 GTTCATCATATATATTCGACAGATTCGTTTGTAAATGAAGGCCACCTCATTTGTCAGAGA 2214
QY 234 rAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAs 253
DB 2215 GAACGTGCCT---CGGGTACTAAATTCAGCTAAGGAGAANTCAAGCACCTGTTCCAAATACC 2271
QY 253 pThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVa 273
DB 2272 TACAGTCAAC-----CAGTATTGTACTTCTTATT 2301
QY 273 lAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTr 293
DB 2302 TGCTCTCATCTTATCTACCGTGACAGCTATCCAGGAATCCACAGTGAAGTGGGTTA 2361
QY 293 pleuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleLeas 313
DB 2362 TGTCGCTATGAAGTTTGCACAGGTCTTTGGTTGCTTTTCTATGTGTACTATCTTTGA 2421
QY 313 pGlnTyrIleAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLe 332
DB 2422 AAGGCTTTGTGCCCTTGTTCGGAAATATCAACAGAGGCCCTTCAGCGCTCGTGTCT 2481
QY 332 uTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352
DB 2482 GGTCTCTATGT-----GTATTTAACTCCATCTTGCCAGGTGCTCTCTCTCTTAC 2535
QY 352 tPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAs 372
DB 2536 TTTTGTGCTTTTGTGCTGCTGCTCAATGCTTTTGTGCTGCTGCTGCTGCTGCTGCTG 2595
QY 372 pArgGluPheLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTr 392
DB 2596 CAGGATGTTCTATAGGATTTGGGAACCTCCAGCTCATCTACTCCAACTATTATAGAACCTG 2655
QY 392 pAsnMetProValHisLysTrpMetIleArgHis-----LeuTyrPhePr 407
DB 2656 GAATGGTGGTCCATGACTGGCTATATATGCTTACAAGGACTTCTCTGCTGCTTTT 2715
QY 407 cCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAl 427
DB 2716 CTCCAAGAGATTC-----AAATCTGCTGCCATGTTAGTGTCTGCTGTATCTGC 2766
QY 427 aLeuPheHisGluLeuCysIleAlaValProCys-----HisIlePh 441
DB 2767 TGTAGTACAGGAATATGCTTGTGCTGTT---TCTTACGCTTTTCTATCCCGTGTCTGT 2823
QY 441 eLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTy 461

Db 2824 CAGTCTCTTCATGTTCTTGGAAATGGCTTC-----ACCTT 2859
Qy 461 rLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSe 481
Db 2860 CATTCGTAATGATAGTCGGAAGAACCCGATTTGGAATGTTCTGATGTGGAC---TTCTC 2915
Qy 481 rIleLeuGlyGlnPrometCysValLeuLeuTyrTyr 493
Db 2916 TTTTCTTGGCAATGGAGTCTTACTCTGCTTTTATTC 2952

RESULT 9

US-08-509-187D-3
; Sequence 3, Application US/08509187D
; Patent No. 5834283
; GENERAL INFORMATION:
; APPLICANT: Chang, Ta-Yuan and Chang, Catherine C.Y.
; TITLE OF INVENTION: ACYL Coenzyme A: Cholesterol Acyltransferase
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/509.187D
; FILING DATE: 31-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lampert Hammitte, Ann
; REGISTRATION NUMBER: 34,858
; REFERENCE/DOCKET NUMBER: DCI-033cpdv
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)42-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1397..3046
US-08-509-187D-3

Alignment Scores:
Pred. No.: 5,21e-35 Length: 4011
Score: 396.00 Matches: 150
Percent Similarity: 44.66% Conservative: 80
Best Local Similarity: 29.13% Mismatches: 191
Query Match: 14.78% Indels: 95
DB: 2 Gaps: 24

US-09-856-018B-16 (1-504) x US-08-509-187D-3 (1-4011)

Qy 33 AsnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaIy 48
Db 1537 GATGCAAGAAAGATAAAGTTTCACAGCAGAGCAGAGGAATTCGAAGCCATTTTATGAA 1596
Qy 48 sAspSerGlySerAspAspSerIleAsnSerAspAspAlaAlaValAsnSerGlnGlnG 68

Db 1597 GGAAGTTGGCAGT-----CACTTTGATGATTTTGTGACCAATCTCATTTGAAA 1644
Qy 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79
Db 1645 GTCAGCATCATATAGATAATGGTGGCGCTCTCACAACCTTTTCTGTCTTCTGAAGAGA 1704
Qy 80 -LysPheAlaTyrArgProSer-----ValPro----- 88
Db 1705 GAAACAACAACCATAGACGGAAGGATTTGAGACACCTCCAGAACGAAGAAAGATTTTAT 1764
Qy 89 -AlaHisArgLysValLysGluSerProLeuSerSerSerThrIlePheArgGlnSerHi 108
Db 1765 TCGAAGCGCTCTCTCTTAGATGAACCTGCTGAAGTGGACACATC----- 1810
Qy 108 salaGlyLeuPheAsnLeuCysIleValValLeuValAla-----ValAsnSerArgLeuI 127
Db 1811 -AGAACAATATATACATGTTTATTGGCCCTCTCATCTCTTTATCTCCTCAGCACACTTGT 1869
Qy 127 eIleGlu-AsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerL 147
Db 1870 AGTAGATTACATTGATGAAGGAAGGTGGTGTGAGT-----TCAGCCT 1914
Qy 147 ysSerLeuArg-AspTrpProLeuPheMetCysCysLeuSerLeuVal----- 162
Db 1915 CCTGTCTTATGCTTTTGGCAAAATTTCTACCGTGTGTTGGACCTGGTGGATCATGTTCT 1974
Qy 163 ValPheProPheAlaAla---PheIleValGluLysLeuAla-----GlnArg 177
Db 1975 GTCTACATTTTTCAGTCTCCCTATTTCTGTTTCAACATTTGGCGCAGCTGGTATGCAAGAG 2034
Qy 178 LysCysIleProGluProValVal-----ValValLeuHisIleIleIleIlehrSer 194
Db 2035 TTCTCATCGCTGATCGGTTCTCTTCCATGGCTTTCTTTTCATGATCTTCCAGATTGG 2094
Qy 195 ThrSerLeuPheTyr-ProValLeuValIleLeuArgCysAspSerAlaPheValSerG1 214
Db 2095 AGTTCTAGGTTTGGACCAACATATGTTGTGTAGCATATACATGCCACCACTCCCG 2154
Qy 214 yValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisTh 234
Db 2155 GTTCATCATATATTCGAGCAGATCGTTTTGTAAAGGCCACTCATTTTGTGACAGAGA 2214
Qy 234 rAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAs 253
Db 2215 GAACGTGCTGCT---CGGGTACTAAATTCAGCTAAGGAGAAATCAAGCAGCTGTTCCAATACC 2271
Qy 253 pThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVa 273
Db 2272 TACAGTCAAC-----CAGTATTCTTCTTCTTAT 2301
Qy 273 lAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTr 293
Db 2302 TGCTCCTACCCCTTATCTACCGTGACAGCTATCCAGGAATCCCACTGTANGATGGGTTA 2361
Qy 293 pLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAs 313
Db 2362 TGTGCTGATGAAGTTTGGACAGGTCTTTGGTGTCTTTTCTATGTGTACTACTCTTTGA 2421
Qy 313 pGlnTyrIleAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLe 332
Db 2422 AAGGCTTTTGCCCTTGTGTTTGGGAATATCAACAGGAGCCCTTCAGCGCTCGTGTCT 2481
Qy 332 uTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352
Db 2482 GGTCTCTATGT-----GTATTTAACTTCCATCTTGCAGGTGTGCTGATCTCTTCTTAC 2535
Qy 352 tPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAs 372
Db 2536 TTTTGTGCTTTTTCGACTGCTGCTCAATGCCCTTGTGAGATGTTTACGCTTGTGGTGA 2595
Qy 372 pArgGluPheTyrLysAspTyrTrpAsnAlaAlaLysThrValGluAspTyrTrpArgMetTr 392
Db 2596 CAGGATGTTCTATAGGATTTGGTGAACCTCCACGTCATCTCCAACTATATAGAACCTG 2655

Db 2422 AAGGCTTTGTGCCCTTGTTCGGAATATCAACACAGAGGCCCTTCAGCGCTCGTGTCT 2481
Qy 332 uTrrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352
Db 2482 GGTCTCTATGT-----GTATTAACTCCATCTTGCAGGTGTGCTGATCTCTCTCTTAC 2535
Qy 352 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAs 372
Db 2536 TTTTCTTTGCTTTTGGCACTGCTGCTCAATGCCCTTCTGCTGAGATGTACGCTTTGGTGA 2595
Qy 372 PArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTr 392
Db 2596 CAGGATGTTCTATAGGATTTGGTGAACCTCCACGTCATCTCACTCACTATATAGAACCTG 2655
Qy 392 pAsnMetProValHisLysTrpMetIleArgHis-----LeuTyrPhePr 407
Db 2656 GAATCGTGTGCTCATGCTGCTATATCTACTATCTTACAGGACTTCTCTGCTGTTT 2715
Qy 407 ocYsLeuArgHisGlyLeuProLysAlaAlaAlaLeuIleAlaPheLeuValSerAl 427
Db 2716 CTCCAAGAGATTC-----AAATCTGCTGCCATGTAGCTGCTTTGCTGTATCTGC 2766
Qy 427 aleuPheHisGluLeuCysIleAlaValProCys-----HisIlePh 441
Db 2767 TGTACTACAGCAATATGCTTGGCTGTT--TGCTTGAGCTTTTCTATCCGCTGCTGT 2823
Qy 441 eLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTy 461
Db 2824 CGTGCCTTCATGTTCTTGTGAATGGCTTTC-----AACTT 2859
Qy 461 rleuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSe 481
Db 2860 CATTCTCAATGATATCGGAAAGCCGATTGGAAATGTTCTGTATGGAC-----TTCTC 2915
Qy 481 rIleLeuGlyGlnProMetCysValLeuLeuTyrTr 493
Db 2916 TTTTCTTGGCAATGAGCTTACTCTGCTTTATTC 2952

RESULT 11

PCT-US93-09704A-3
; Sequence 3, Application PC/TUS9309704A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL
; TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)
; NUMBER OF SEQUENCES: 9
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09704A
; FILING DATE: October 12, 1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. SER. NO. 959,950
; FILING DATE: October 14, 1992
; APPLICATION NUMBER: U.S. SER. NO. 121,057
; FILING DATE: September 10, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LAMPORT HAMMITTE, ANN.
; REGISTRATION NUMBER: 34,858
; REFERENCE/DOCKET NUMBER: DCI-033CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-2700
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: CDNA
PCT-US93-09704A-3
Alignment Scores:
Pred. No.: 5,21e-35 Length: 4011
Score: 396.00 Matches: 150
Percent Similarity: 44.66% Conservative: 80
Best Local Similarity: 29.13% Mismatches: 191
Query Match: 14.78% Indels: 95
DB: 5 Gaps: 24
US-09-856-018B-16 (1-504) x PCT-US93-09704A-3 (1-4011)
Qy 33 AsnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaLys 48
Db 1537 GATACAAAGAGATAAAGATTTCACAGCAGAGAGCAGAGGAATTAAGCCATTATTTATGAA 1596
Qy 48 sAspSerGlySerAspAspSerIleAsnSerAspAspAlaAlaValAsnSerGlnGlnI 68
Db 1597 GGAAGTTGGCAGT-----CACTTTGATGATTTTGTGACCAATCTCATTTGAANA 1644
Qy 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79
Db 1645 GTCACGATCATTAGATAGATAATGGTGGTGGCTCTCTACACCTTTTCTGTCTTGAAGGAGA 1704
Qy 80 -LysPheAlaTyrArgProSer-----ValPro----- 88
Db 1705 GAAACCAACCATAGACGCAAGGATTTGAGAGCACCTCCAGAACAGAAAGATTTTAT 1764
Qy 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHi 108
Db 1765 TGCAAGGCGCTCTCTCTTAGATGAACCTGCTTGAAGTGACCAATC----- 1810
Qy 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeuI 127
Db 1811 -AGAACATATATACATGTTTATTTGCGCTCTCTCATTTCTTTATCTCAGCACACTGT 1869
Qy 127 eIleGlu-AsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerL 147
Db 1870 AGTAGATTACATTGATGAAGGAAGCTGCTGTGAGT-----TCAGCCT 1914
Qy 147 ySerLeuArg-AspTrpProLeuPheMetCysCysLeuSerLeuVal----- 162
Db 1915 CTTGCTTATGCTTTTGGCAAAATTTCTTACCGTCTTTGGACCTGGTGATCATCTTCTCT 1974
Qy 163 ValPhePropheAlaAla---PheIleValGluLysLeuAla-----GlnArg 177
Db 1975 GTCTACATTTTCAGTTCCTCTATTTTCTGTTCAACATTGGCGCAGCTGCTATAGCAAGAG 2034
Qy 178 LysCysIleProGluProValVal-----ValValLeuHisIleIleIleThrSer 194
Db 2035 TTCTCATCCGCTGATCCGTTCTCTCTTCCATGGCTTTCTTTTCATGATGATCTTCCAGATTGG 2094
Qy 195 ThrSerLeuPheTyr-ProValLeuValIleLeuArgCysAspSerAlaPheValSerGl 214
Db 2095 AGTCTAGTGTGGACCAACATATGTTGTGTAGCATATACATGCCACAGCTTCCG 2154
Qy 214 yValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisTh 234
Db 2155 GTTCATCATTTATTCGAGCAGATTCGTTTGTGTAATGAAGGCCCACTCTTTTGTCCAGAGA 2214
Qy 234 rAsnTyrAspMetArgAlaLeuThrLysLeuValGlyGlyGluAlaLeu---LeuAs 253
Db 2215 GAACGTGCTCT---CGGGTACTAAATTCAGCTAAAGGAGAAATCAAGCACTGTTTCCAATACC 2271
Qy 253 pThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVa 273
Db 2272 TACAGTCAAC-----CAGTATTGCTACTTCTTAT 2301
Qy 273 lAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTr 293
Db 2302 TGCTCTACCTTATCTACCGTGACAGCTATCCAGGAATCCACCTGTAAAGATGGGGTTA 2361

QY 293 pLeuPheArgGlnLeuValLysLeuLeuLeuPheThrGlyValMetGlyPheLeuLeuLeu 313
Db 2362 TGTGCGTATGAAGTTTGCACAGCTTTGGTGTCTTTCTTCTATGACTACATCTTTGA 2421
QY 313 pGlnTyrIleAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLe 332
Db 2422 AAGCGTTTGTGCGCCCTTTGTCGGAATATCAAAACAGGAGCCCTTCAGGCGCTCGTGTCT 2481
QY 332 uTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352
Db 2482 GGTCTCTATGT-----GTATTTAACTCCATCTGTGCCAGGTGCTGATCTCTCTCTTAC 2535
QY 352 tPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAs 372
Db 2536 TTTTGTGCGCTTTTGCACCTGCTGCGTCAATGCGCTTGTGCTGAGATGTTACGCTTTGGTGA 2595
QY 372 pArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTr 392
Db 2596 CAGGATGTTCTATAGGATGGTGGAACTCCAGTCATCTCAACTATATAGAACCTG 2655
QY 392 pAsnMetProValHisLysTrpMetIleArgHis-----LeuTyrPhePr 407
Db 2656 GAATGGTGGTCCACTGCTGGCTATATCTATGCTTACAAGGACTTTCTCTGGTTTTT 2715
QY 407 oCysLeuArgHisGlyLeuProLysAlaLaLeuLeuLeuAlaPheLeuValSerAl 427
Db 2716 CTCAAAGAGATTC-----AAATCTGCTGCATGTTAGCTGTCTTGTCTATCTGC 2766
QY 427 aLeuPheHisGluLeuCysIleAlaValProCys-----HisIlePh 441
Db 2767 TGTAGTACAGATATGCTGCTGCTGTT---TGCTGAGCTTTTCTATCCCGTGTGTT 2823
QY 441 eLysLeuTrpAlaPheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTy 461
Db 2824 CGTGCCTTCATGCTTCTTTGGAATGCTTC-----AACIT 2859
QY 461 rLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSe 481
Db 2860 CATTTGCAATGATAGTCGGAAGGCGGATTTGGAAATGTTCTGATGTGGAC-----TTCTC 2915
QY 481 rIleLeuGlyGlnProMetCysValLeuLeuTyrTyr 493
Db 2916 TTTTCTGGCAATGAGTCTTACTCTGCTTTTATTC 2952

RESULT 12

US-09-165-042-4
; Sequence 4, Application US/09165042
; Patent No. 6100077
; GENERAL INFORMATION:
; APPLICANT: Sturley, Stephen L.
; APPLICANT: Oskers, Peter
; TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL
; TITLE OF INVENTION: ACYLTRANSFERASE
; FILE REFERENCE: 0575/56331
; CURRENT APPLICATION NUMBER: US/09/165,042
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Yeast
US-09-165-042-4

Alignment Scores:

Pred. No.:	8, 44e-35-	Length:	2040
Score:	350.00	Matches:	143
Percent Similarity:	42.73%	Conservative:	92
Best Local Similarity:	26.00%	Mismatches:	202
Query Match:	14.55%	Indels:	113
DB:	3	Gaps:	26

US-09-856-018b-16 (1-504) x US-09-165-042-4 (1-2040)

QY 16 HisSerSerLeuArgArgArgProSerAlaThrSer-----ThrAlaGlyLeuPheAsnSe 34
Db 48 CACCATGGAGCGCGCGCGCGCTCTGCTGCTGCAGAGACAGAGGCGTG-----GG 101
QY 34 rProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGlySerAspAs 54
Db 102 AGGGAGCGGGAGCGCAACCCCTGTGGAGAT-----GGAAACACTGA 143
QY 54 pSerIleAsnSerAspAsp-----AlaAlaValAsnSerGI 66
Db 144 GAGCGACAGAGCGCGCGCGCTTGGTACAATGGACCGACACATGGAGGCTGTGAAGCACA 203
QY 66 nGlnGlnAsnGluLysGlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSe 86
Db 204 ATTGCTGGAGCAAGCGCGAGCAACTGAGGAGGCTGCTGGATCGGCCCATCGGGAGGC 263
QY 86 rValProAlaHis-----Ar 91
Db 264 TATACAATCTTACCCATCACAAGACAAACCTCTGCCCCACCTCCCGAGGTTCTCTTGAG 323
QY 91 gLysValLysGluSerProLeuSerSerAspThrIlePhe-----ArgGlnSer----- 107
Db 324 CAGGACCCAGGAGCCATCCCTGGGAAACAGAAAGTTTTCATCATCCGCAAGTCCCTGCT 383
QY 108 -----HisAlaGlyLeuPheAsnLe 114
Db 384 TGATGAGCTGATGGAGGTGCAGCATTTCCGACCATCTACCACATGTTTCATCGCTGGCCT 443
QY 114 uCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsn-----LeuMe 132
Db 444 GTGTGTCTTCATCATCAGCACCCCTGGCCATCGACTTCATTGATGAGGCGAGGCTGTGCT 503
QY 132 tLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTr 152
Db 504 GGAGTTGAGCTACTGATCTTCTCAGC---TTCGGACAGCTGCCATTGGCGCTGGTGACCTG 560
QY 152 p---ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleVa 171
Db 561 GGTGCCCATGTTT-----CTGTCCACCCTGTGTGGCGCGTACCAGGCCCTA----- 606
QY 171 lGluLysLeuAlaGlnArgLysCysIleProGluPro-----ValValVa 186
Db 607 ----CGGCTGTGGCGCAGGGGACCTTGGACGAGCGGCGCTGGGCTGGGCTGTGCTGTG 662
QY 186 lValLeuHisIleIleIleThrSerThrSerLeuPheTyrProValLeuValIleLeuAr 206
Db 663 AGCGCGCCACGCGGTGCTGTGCGCG-----CTGCGCGTCCACGTGCCCGCTGGA 713
QY 206 gCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSerCysValValTrpLe 226
Db 714 GCATCAGCTCCCGCGCGCTCC-----CGTTGTGCTCTGTGCTCT 752
QY 226 uLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGl 246
Db 753 CGAGCAGGTAGGTTCCTGATGAAAAGTACTCTCTCTCCTGAGA----- 795
QY 246 uLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnVal----- 263
Db 796 -----GAGGCTGTGCTGGGACCTTCTGTGCCACAGGAGGTGAGGGATCCAGGCCCC 848
QY 264 -SerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTy 283
Db 849 CAGTTTCTCCAGCTACTCTACTTCTCTTCTGCCCCAACACATCATCTACAGGAGACTTA 908
QY 283 rProArgThrProTyrIleArgLysGlyTrpLeu-----PheArgGlnLeuVally 300
Db 909 CCTAGGACGCCCTATGTGAGGTGGAATATGTGGCCAAAGAACTTTGCCAGGCCCTGGG 968
QY 300 sLeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleVa 320
Db 969 ATGTGTGCTCTATGCC-----TGCTTCATCTGCGCGCGCTCTGTGTTCTCTGCTCT 1019

```

Qy 320 IclnAsn---SerGlnHisProLeuLysGlyAsnLeuLeuYrAlaThrGluArgValle 339
Db 1020 TGCCAAACATGAGCCGAGCCCTCCAGCCCGTCCCTGGCTCTCT-----ATCCT 1073
Qy 339 uLysLeuSerValProAsnLeuYrValTrpLeuLysMetPheTrpCysPheHisLe 359
Db 1074 GCATCCACGTTGCCAGCATCTTCATCTGCTGCTATCTCTTCTTCTCTCCATG 1133
Qy 359 uTrpLeuAsnLeuLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTrpLysAspTr 379
Db 1134 CTGGCTCAACGCTTTGCCGAGATCTACGATTTGGACACAGGATGTTCTACCGGACTG 1193
Qy 379 pTrpAsnAlaLysThrValGluAspTrpArgMetTrpAsnMetProValHisLysTr 399
Db 1194 GTGGAACTCAACGCTCTCTCCACTACTACCGCACTTGGAAACGTTGGTGGTCCATGACTG 1253
Qy 399 pMetIleArgHisLeuYrPheProCysLeuArg-----HisGlyLeuProLysAlaAl 417
Db 1254 GCTGTACAGTACGTTATCAGGATGGGCTCGGCTCTTGTGTCGCCGCGGAGGGT 1313
Qy 417 aAlaLeuLeuLeuAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValPr 437
Db 1314 AGCCATGCTGGTGTGTTCTCTCGGCTCGGCACTGGCCCATGAGTAT-----ATCTT 1364
Qy 437 ocYsHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValle 457
Db 1365 CTGCTTCCTCTCTG-----GGGTTCTTCTATCCGTCATGCTGCTACT 1406
Qy 457 u-----IleThrAsnTrpLeuGlnAsnLysPheArgAsnSerMetVa 471
Db 1407 CTTCCTTGTGATGGAGGATTTGAATCTCATGATGATGATGATGATGATGATGATGATG 1466
Qy 471 IclnAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLe 491
Db 1467 ATGGAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1523
Qy 491 uTrpTrpHisAspLeuMetAsnArgLys 500
Db 1524 GTATGCCAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551

```

RESULT 13

```

US-08-121-057-2
; Sequence 2, Application US/08121057
; Patent No. 5484727
; GENERAL INFORMATION:
; APPLICANT: CHANG, TA-YUAN
; TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL
; TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121.057
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: LAMFORT HAMMITE, ANN.
; REGISTRATION NUMBER: 34,858
; REFERENCE/DOCKET NUMBER: DCI-033CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-2700
; TELEFAX: (617) 227-5941

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-121-057-2

```

```

Alignment Scores:
Pred. NO.: 1.08e-32 Length: 4079
Score: 376.00 Matches: 151
Percent Similarity: 42.94% Conservative: 77
Best Local Similarity: 28.44% Mismatches: 175
Query Match: 14.03% Indels: 130
DB: 1 Gaps: 26

```

US-09-856-018B-16 (1-504) x US-08-121-057-2 (1-4079)

```

Qy 33 AsnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaLys 48
Db 1608 CATGCAAGAGAGATAAAGTTTGACAGCAGAGCAGAGGAATTTGAAGCCATTTTATGAA 1667
Qy 48 aspSerGlySerAspAspSerIleAsnSerAspAspAlaAlaValAsnSerGlnGlnCl 68
Db 1668 GGAAGTTGGCAGT-----CAGTTGATGATTTTGTGACCAATCTCATTTGAAA 1715
Qy 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu-----79
Db 1716 GTCACCATCATAGATATATGTTGGTGGCTCTCACACCTTTCTGTCTTGAAGGAGA 1775
Qy 80 -LysPheAlaTrpArgProSer-----ValPro-----88
Db 1776 GAAACCAACATAGACGAGGAGGATTTGAGAGCAGCTCCAGAACAGGAAGATTTTAT 1835
Qy 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHi 108
Db 1836 TGCAAGGCGCTCTCTTAGTAGTAAGTCTGTTGAAGTGGACACATC-----1881
Qy 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeu11 127
Db 1882 -AGAACATATATACATGTTTATTTGCCCTCTCTCATTTCTTTATCCTCAGCACACTGT 1940
Qy 127 eLleGlu-AsnLeuMetLysTrpLysLys-----139
Db 1941 AGTAGATTACATTTGATGAAGGAGGCTGCTGCTTCAAGTTACGCTCTCTCTTCCAT 2000
Qy 140 ----SerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPhe-----155
Db 2001 TTTTGGCAAAATTC-----CTACCGTTGTTGGACCTG 2033
Qy 156 ----MetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGluLysL 174
Db 2034 GTGGATCATGTTCTCTACATTTTTCAGTTCCTCTCTTCTTCTTCAACATGGCGGAC 2093
Qy 174 euAla---GlnArgLysCysIleProGluProValValValLeuHisIleIleIle 193
Db 2094 TGGCTATAGCAAGAGCTTCTCATCCGCTGATCGCTCTCTCTTCCATGGCTTCTTTT 2153
Qy 193 hrSerThrSerLeuPheTrpProValLeu-----202
Db 2154 GATCTTCCAGATTGGAGTTCTAGGTTTGGACCAACATATGTTGTTAGCATATCTCTGC 2213
Qy 203 -----ValIleLeuArgCysAspSerAlaPheValSerGlyValThrL 217
Db 2214 CACAGCTTCCCGGTTTCATCATTTATTCGAGCAGAT---CGTTTCTGA-----2257
Qy 217 euMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTrp 237
Db 2258 -----ATGAGGCCCACTCAATTTCTCAGAGAGAACGTCG 2291
Qy 237 spMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAspThrLeuA 256

```

Db 2292 CT---CGGGTACTAATTCAGCTAAG-CAGAAATCAAGCAGCTGTTCCAAATCCTACAGTCA 2347
QY 256 snMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProT 276
Db 2348 AC-----CAGTATTGCTACTTCTTATTTGCTCCTA 2377
QY 276 hrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheA 296
Db 2378 CCCTATCTACCGTGACAGCTATCCAGGAATCCCACTGAAGATGGGTATGTTGCTA 2437
QY 296 rgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleaspGlnTyrI 316
Db 2438 TGAAGTTGGACAGCTGTTGGTGTCTTTCTATGTGCTACTATTTGAAAGGCTTT 2497
QY 316 leAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaT 335
Db 2498 GTGCCCCCTGTTTCGGAATATCAACAGGAGCCCTTCAGCGCTCGTGTCTGCTCTAT 2557
QY 335 hrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrC 355
Db 2558 GT-----GTATTTAACTCCATCTTGCAGGTGCTGCTGATTTCTTCTTACTTTTGTG 2611
QY 355 ysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluP 375
Db 2612 CCTTTTGCAGCTGCTGCTCAATCCCTTGTGAGATGTACGCTTTGGTGACAGATG 2671
QY 375 heTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetP 395
Db 2672 TCTATAAGGATGTGTGAACCTCCACGTCATACCTCAACTATTATAGAACCCTGGAATCGG 2731
QY 395 roValHisLysTrpMetIleArgHis-----LeuTyrPheProCysLeuA 410
Db 2732 TGGTCCATGACTGGCTATATTACTATCTTACAGGACTTTCTGCTGTTTCTTCCAGA 2791
QY 410 rgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheH 430
Db 2792 GATTC-----AAATCTGCTGCCAATGTAGCTGCTTTGCTGTATCTGCTGATGAC 2842
QY 430 lsGluLeuCysIleAlaValProCys-----HisIlePheLysLeuT 444
Db 2843 ACGAATATGCTTGGCTGTT---TGCTTGAGCTTTTCTATCCGCTGCTGCTGCTGCT 2899
QY 444 rPalapheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnA 464
Db 2900 TCATGTTCTTTGGAATGGCTTTC-----AACTTCAATTCCTCA 2935
QY 464 snLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuG 484
Db 2936 ATGATAGTCGGAAACCCGATTTGGATGTTCTGATGTGGAC-----TTCTCTTTTCTTG 2991
QY 484 lyGlnProMetCysValLeuLeuTyrTyr 493
Db 2992 GGCAATGGAGTCTTACTCTGTTTATTC 3020

RESULT 14

US-08-509-187D-2
; Sequence 2, Application US/08509187D
; Patent No. 5834283
; GENERAL INFORMATION:
; APPLICANT: Chang, Ta-Yuan and Chang, Catherine C.Y.
; TITLE OF INVENTION: ACYL Coenzyme A: Cholesterol Acyltransferase
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/509,187D
; FILING DATE: 31-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lamport Hammitte, Ann
; REGISTRATION NUMBER: 34,858
; REFERENCE/DOCKET NUMBER: DCI-033cpdv
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-509-187D-2

Alignment Scores:
Pred. No.: 1.08e-32 Length: 4079
Score: 376.00 Matches: 151
Percent Similarity: 42.94% Conservative: 77
Best Local Similarity: 28.44% Mismatches: 175
Query Match: 14.03% Indels: 130
DB: 2 Gaps: 26

US-09-856-018B-16 (1-504) x US-08-509-187D-2 (1-4079)

QY 33 AsnSerProGluThrThrAspSerSerGly-AspAspLeu-----AlaLy 48
Db 1608 GATAGCAAGCAAGATAAAGTTGACAGCAGGAGCAAGAAATTGAAGCCATTTTATGAA 1667
QY 48 saspSerGlySerAspAspSerIleAsnSerAspAlaAlaValAsnSerGlnGlnGl 68
Db 1668 GGAAGTTGGCAGT-----CATTGTGATGATTGTTGACCACATCTCATTGAAAA 1715
QY 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu-----79
Db 1716 GTGACGATCATATTAGATAATAGTGGTGGCTCTCACAACTTTCTGTTCTTGAAGCAGA 1775
QY 80 -LysPheAlaTyrArgProSer-----ValPro-----88
Db 1776 GAAACCAACCATAGCGGAGGAGATTGAGAGCACCTCCAGAACAAAGAAAGATTTTAT 1835
QY 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerH 108
Db 1836 TGCAAGCGGCTCTCTCTTAGTGAAGTGGACCATC-----1881
QY 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeuI 127
Db 1882 -AGAACCAATATATACATGTTTATGGCCCTCTCTCTTCTTATCTCTCAGCACACTGT 1940
QY 127 eileGlu-AsnLeuMetLysTyrGlyTrpLeuIleLys-----139
Db 1941 AGTAGATTACATTGATGAAGGAGGCTGGTGTGCTTACGCTCCTGCTTATGCA 2000
QY 140 -----SerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPhe-----155
Db 2001 TTTTGGCAAAATTC-----CTACCGTGTCTTGGACCTG 2033
QY 156 -----MetCysCysLeuSerLeuValValPheProPheAlaAlaIleValLysL 174
Db 2034 GTGGATCATGTTCTGCTACATTTTTCAGTTCCCTATTTTCTGTTTCAACATTTGGCGCAC 2093
QY 174 euAla---GlnArgLysCysIleProGluProValValValLeuHisIleIleIle 193
Db 2094 TGGCTATAGCAGAGATTCTCATCCGCTGATCCGCTTCTCTCTTCCATGGCTTTCTTTTCAT 2153

QY 193 hrSerThrSerLeuPheTyrProValLeu----- 202
Db 2154 GATCTTCCAGATTGGAGTCTTAGTGGTGGACCAACATATGTTGTTAGCATATCCTGC 2213
QY 203 -----ValIleuArgCysAspSerAlaPheValSerGlyValThrL 217
Db 2214 CACCAGCTTCCGGTTCATCATTAATTCGACGAGAT---CGTTTGTA----- 2257
QY 217 euMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrA 237
Db 2258 -----ATGAAGGCCCACTCATTTGTCAGAGAGACGTGC 2291
QY 237 spMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAspThrLeuA 256
Db 2292 CT---CGGTACTAATTCAGCTAAG-GAGAAATCAACGACTGTTCCAATACCTACAGTCA 2347
QY 256 snMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProt 276
Db 2348 AC-----CAGTATTGCTACTCTTATTGCTCTCTA 2377
QY 276 hrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheA 296
Db 2378 CCCATTATCCGGTACAGCTATCCAGGAATCCCACTGTAAGATGGGTTATGCTA 2437
QY 296 rgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheIleIleAspGlnTyrI 316
Db 2438 TGAAGTTTGCACAGCTCTTGGTGTCTTTTCTATGTGTACTACATCTTTGAAAGGCTTT 2497
QY 316 leAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaT 335
Db 2498 GTGCCCCCTTGTTCGGAATATCAACAGGAGCCCTTCAGGCGTCGTTCTGTCCTAT 2557
QY 335 hrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrC 355
Db 2558 GT-----GTATTTAACTCCATCTGCGAGTGTGCTGATCTCTCTTACTTTTGG 2611
QY 355 ysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluP 375
Db 2612 CCTTTTTCACACTGTGCTCAATCCCTTGTGCTGAGATGTTACGCTTGTGTCACAGGATGT 2671
QY 375 heTyrLysAspTyrTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetP 395
Db 2672 TCTATAGGATTGTGGAACTCCAGCTCACTCCAACTATTATAGAACCTGGGATGG 2731
QY 395 roValHisLysTrpMetIleArgHis-----LeuTyrPheProCysLeuA 410
Db 2732 TGGTCCATGACTGCTATATTACTATGCTTACAAGGACTTCTCTGTTTTCCTCCAAGA 2791
QY 410 rgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheH 430
Db 2792 GATTC-----AAATCTGCTGCCATGTAGTGTCTTTGCTGTATCTGCTAGTAC 2842
QY 430 isGluLeuCysIleAlaValProCys-----HisIlePheLysLeuT 444
Db 2843 ACGAATATGCTTGGCTGCTTT---TGCCTGAGCTTTTCTATCCCGTGTGCTGCTCT 2899
QY 444 rpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnA 464
Db 2900 TCATGTTCTTTGGAATGGCTTC-----AATCTCATTTGCTCA 2935
QY 464 snLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuG 484
Db 2936 ATGATAGTCGGAAGAACCCGATTTGGAATGTTCTGTGTGGAC---TTCCTTTTCTCTG 2991
QY 484 lyGlnProMetCysValLeuLeuTyrTyr 493
Db 2992 GGCAATGGAGCTTACTCTGCTTTTATTC 3020

RESULT 15

US-09-121-396-2

: Sequence 2: Application US/09121396

: Patent No. 5968749

GENERAL INFORMATION:
APPLICANT: CHANG, TA-YUAN
TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL
TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/121,396
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/509,187
FILING DATE: 07/31/95
ATTORNEY/AGENT INFORMATION:
NAME: LAMPORT HAMMITTE, ANN.
REGISTRATION NUMBER: 34,858
REFERENCE/DOCKET NUMBER: DCI-033CPDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-2700
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4079 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-121-396-2

Alignment Scores:
Pred. No.: 1.08e-32 Length: 4079
Score: 376.00 Matches: 151
Percent Similarity: 42.94% Conservative: 77
Best Local Similarity: 28.44% Mismatches: 175
Query Match: 14.03% Indels: 130
DB: 2 Gaps: 26

US-09-856-018B-16 (1-504) x US-09-121-396-2 (1-4079)
QY 33 AsnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaLys 48
Db 1608 GATAGCAAGAGAGATAAGTTGACAGCAGCAGAGAGGATTTGAAGCCATTTTATGAA 1667
QY 48 sAspSerGlySerAspSerIleAsnSerAspAlaAlaValAsnSerGlnGlnI 68
Db 1668 GGAAGTTGGCAGT-----CACTTTGATGATTTTGTGACCAATCTCATTTGAAA 1715
QY 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79
Db 1716 GTCAGCATCATATTAGATAATGTTGGTGGCTCTCCACACCTTTTCTGTTCTTGAAGGAGA 1775
QY 80 -LysPheAlaTyrArgProSer-----ValPro----- 88
Db 1776 GAAAACAACCATAGACGAGGAGATTGAGAGCACCTCCAGACAAGAAAGATTTTAT 1835
QY 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerH 108
Db 1836 TGCAAGCGCTCTCTCTTAGATGAACCTTGAAGTGGACCATC----- 1881
QY 108 sAlaGlyLysPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeuI 127
Db 127

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2003, 19:43:31 ; Search time 1672 Seconds
(without alignments)
4881.893 Million cell updates/sec

Title: US-09-856-018b-16
Perfect score: 2680
Sequence: 1 MAISDEPESVATLHNSLR.....QPMCVLLYHDLNRKGLD 504

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xl
-Q/cgn2_1/spool/US09856018/runat_20022003_091206_8351/app_query.fasta_1.647
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcg -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856018 @CGN_1_1_2874_@runat_20022003_091206_8351 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1321	49.3	1572	11	AY105372	Zea mays
C	1187	44.3	961	14	BQ510367	AY105372 Zea mays
3	1107	41.3	1181	11	AY110660	AY110660 Zea mays
4	1038.5	38.8	729	13	BI422326	BI422326 EST332932
C	952	35.5	586	14	BQ148998	BQ148998 NF086D09F
6	885	33.0	855	10	AW349274	AW349274 GM210004B
7	880	32.8	572	13	BM309699	BM309699 sak65f03.
8	871	32.5	741	14	BQ989822	BQ989822 OGG20J23.
C	854	31.9	729	13	BQ322102	BQ322102 BQ322102
10	823.5	30.7	572	14	BQ124305	BQ124305 EST609881
11	800	29.9	606	13	BQ316561	BQ316561 BQ316561
12	784	29.3	539	13	BI422212	BI422212 EST532878
13	761	28.4	540	9	AJ470192	AJ470192 AJ470192
14	732.5	27.3	561	10	AW035727	AW035727 EST281881
15	729	27.2	433	9	AI441040	AI441040 sa58f02.y
16	707	26.4	626	10	AV926912	AV926912 AV926912
17	691	25.8	862	12	BG321213	BG321213 Zm04_05G0
18	660	24.6	682	10	BE247899	BE247899 NF038D11D
19	657	24.5	577	10	AW586836	AW586836 EST318459
20	652	24.3	763	13	BI308446	BI308446 EST529856
21	643	24.0	655	14	BQ988311	BQ988311 OGG19C15.
22	635	23.7	480	9	AJ470191	AJ470191 AJ470191
23	632	23.6	480	9	AJ470190	AJ470190 AJ470190
24	626.5	23.4	459	14	BQ123670	BQ123670 EST609246
25	620	23.1	634	12	BF634363	BF634363 NF059D06D
26	608	22.7	1043	14	BM919422	BM919422 AGENCOURT
27	601.5	22.4	673	14	BQ988554	BQ988554 OGG19N20.
28	597.5	22.3	1064	13	BM476315	BM476315 AGENCOURT
29	597	22.3	508	14	BU009402	BU009402 QGJ10F20.
30	579	21.6	555	13	BQ268713	BQ268713 BQ268713
31	572	21.3	685	14	BQ861203	BQ861203 QCC17N02.
32	568.5	21.2	629	9	AA042298	AA042298 24635 CD4
33	559.5	20.9	619	14	BM739263	BM739263 K-EST00008
34	559	20.9	515	10	AV925760	AV925760 AV925760
35	558.5	20.8	968	14	BQ670343	BQ670343 AGENCOURT
36	555.5	20.7	627	10	AV956927	AV956927 AV956927
37	549	20.5	656	14	BQ862755	BQ862755 QCC21P16.
38	543	20.3	819	13	BI225365	BI225365 G02950108
39	540.5	20.2	661	14	BQ042483	BQ042483 UI-M-EMO-
40	539	20.1	655	10	AW775077	AW775077 EST334228
41	529.5	19.8	585	12	BQ078967	BQ078967 H3034B09-
42	524	19.6	574	12	BF003479	BF003479 EST431977
43	516.5	19.3	570	14	BM829896	BM829896 K-EST0103
C	516	19.3	617	10	AV930897	AV930897 AV930897
C	501	18.7	857	9	AL520752	AL520752 AL520752

ALIGNMENTS

RESULT 1

AY105372
LOCUS AY105372
DEFINITION Zea mays
ACCESSION AY105372
VERSION AY105372.1
KEYWORDS HTc
SOURCE Zea mays.
ORGANISM Zea mays

1572 bp
PC0148220 mRNA sequence.
GI:21208450
1 (bases 1 to 1572)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

1572 bp
PC0148220 mRNA sequence.
GI:21208450
1 (bases 1 to 1572)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

1572 bp
PC0148220 mRNA sequence.
GI:21208450
1 (bases 1 to 1572)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 1572)
 AUTHORS Coe,E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 FEATURES Location/Qualifiers
 1..1572
 /organism="Zea mays"
 /db_xref="MaizeDB:630748"
 /db_xref="taxon:4577"
 /clone_lib="PCOI48220"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 438 a 294 c 285 g 555 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.79e-133 Length: 1572
 Score: .1321.00 Matches: 235
 Percent Similarity: 66.90% Conservative: 54
 Best Local Similarity: 54.40% Mismatches: 72
 Query Match: 49.29% Indels: 71
 DB: 11 Gaps: 1

US-09-856-018b-16 (1-504) x AY105372 (1-1572)

Qy 142 PheTrpPheSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerLeu 161
 Db 1 TTTTGGTTAATGCTACATCATGCGAGACTGGCCACCTGAATGTGTGGCTTAGTCTA 60
 Qy 162 ValValPheProPheAlaAlaPheValLeuGluLysLeuAlaGlnArgLysCysLeuPro 181
 Db 61 CCCATATTTCCCTTGGTGCACTTTCAGTCGAAAGATTGGCATTCAACAATCTCGTTAGT 120
 Qy 182 GluProValValValLeuHisIleIleIleIleIleIleIleIleIleIleIleIleIle 201
 Db 121 GATCGTCTACTACCTGTTTTCACATCTCTTTTACACATTTTGAATTTGATATATCCAGT 180
 Qy 202 LeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSer 221
 Db 181 CTCGGAATCTTAAGTGTGATCTCCAGTTTATCAGGCTTTGTGTTGATGTTTATTGCC 240
 Qy 222 CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeu 241
 Db 241 TGCATTTGTTGGCTGAAGCTGTATCTTTTGCACATACAAACCATGATATAGAAACATG 300
 Qy 242 ThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyr 261
 Db 301 ATCACAAGCGCAGAGAGTTGATTAATGAAGTACCGCGGCTGGCATAGATAATTTACAA 360
 Qy 262 AsnValSerPheLysSerLeuAlaTrpPheLeuValAlaProThrLeuCysTyrGlnPro 281
 Db 361 GCTCAACTCTGGGAGTCTACATCTTCATGATGGCTCCGACACTCTGTATCAGCCA 420
 Qy 282 SerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuLysLeu 301
 Db 421 AGTTATCTCGAACACCTTATGTTAGAAAGTTGGCTGGTCCGTCAGTATTCTATAC 480
 Qy 302 IleIlePheThrGlyValMetGlyPheIleIleLeuAspGlnTyrIleAsnProIleValGln 321
 Db 481 TTGATATTACTGGTCTCCAAGATTCATTATTGAGCAATACATAAATCCATTCTGTG 540
 Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLeu 341

Db 541 AACTCTCAACATCCATTTGATGGGAGGATTACTGAATGCTGTAGAGACTGTTTTGAAGCTC 600
 Qy 342 SerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeu 361
 Db 601 TCATTACCAATGCTACTCTGCTGGCTTGCATGTTTATTGCTTTTCCATCTGTGGTTA 560
 Qy 362 AsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsn 381
 Db 661 AACATACTGCTGAGATTCTCGATTGGTGAGCCGAGAAATCTACAAAGACTGGTGAAT 720
 Qy 382 AlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIle 401
 Db 721 GCAAGACAAATTTGATGAGTACTGGAGAAATGGAATGCTGCTCATAAATGGATTGTT 780
 Qy 402 ArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIle 421
 Db 781 CGTCATATATATTTCCCTGTCATGCGAATGATATATCAAGGAAGTGTCTGTTTTATA 840
 Qy 422 AlaPheLeuValSerAlaLeu-Phe- 429
 Db 841 TCGTTCTTTGTTCTGCTACTTCTCATGAGGTAACCTATTATTCTTTTCACTCTTCATCT 900
 Qy 429 ----- 429
 Db 901 GCATATATTAATATATATAGTCTCTATTATTTCAAAATGTGTCTTTCGAGTTTCGACATGCT 960
 Qy 429 ----- 429
 Db 961 TTTGTTCAAACTTACCAGCTGTAGATTACTTGGATGAAGTGTCTATATATAAATCAATA 1020
 Qy 430 -----HisG1 431
 Db 1021 TTTCACAATCCAGTCCCTTTCGAGAAATTTATGATACATTTTGTGTCATTTGTACACCA 1080
 Qy 431 uLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPh 451
 Db 1081 GTATGCTTTGCGATTGCTCCCATATCAAGTCTCTGGGCTTTCTAGGAATCATGCT 1140
 Qy 451 eGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetVa 471
 Db 1141 TCAGATTTCCCTCATCATATTGACATCATCTCAAAATAAATTCAGTGACACAAATGGT 1200
 Qy 471 lGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLe 491
 Db 1201 TGGCAATATGATCTTTTGTGTTTTTCTGATATACGGCAGCAATGTGTGTCTATT 1260
 Qy 491 uTyrTyrHisAspLeuMetAsnArgLysGlyLys 502
 Db 1261 GTATTACCATGATGTGATGAACCGGACTGAGAAG 1294

RESULT 2
 BQ510367/c

LOCUS BQ510367 961 bp mRNA linear EST 22-JUL-2002
 DEFINITION EST617782 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ76 3' end, mRNA sequence.

ACCESSION BQ510367
 VERSION BQ510367.2 GI:21926064
 KEYWORDS EST.

SOURCE Solanum tuberosum

ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 961)

AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.

TITLE Generation of a set of potato cDNA clones for microarray analyses

JOURNAL Unpublished (2002)

COMMENT On Jun 10, 2002 this sequence version replaced gi:21369236. Other ESTs: EST617781
 Contact: Robin Buell


```

Db 72 GCTGCATATGGAATATGTCGATCTCTGAGAAATGAAAGATCCA-----ACCTTT 122
Qy 266 LysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArg 285
Db 123 AAAAGCTAGTGTACTTCATGTTGGCCCAACACTTGTACAGCACTATCTCTCAA 182
Qy 286 ThrProTyrIleArgLysGlyTyrPheLeuValLysLeuValLysLeuPheThr 305
Db 183 ACTACATGTATGAAAGGTTGGTGACCCCACTATAAAGTGGTGGNNNNA 242
Qy 306 GlyValMetGlyPheIleAspGlnTyrIleAsnProIleValGlnAnSerGlnHis 325
Db 243 GGCTGTATGGCTTCATATGAGCAATATATAACCAATGTGAAGATTCACAAAT 302
Qy 326 ProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsn 345
Db 303 CCAGTAAAGGGAANNNGAATGCTATAGAAAGAGTCTTAAACTCTCAGTCCCAACA 362
Qy 346 LeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAla 365
Db 363 TTATATGTATGGCTTTCATGCTTCTATTCNNNNNCATTTATGGCTGAACATTTAGCT 422
Qy 366 GluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrVal 385
Db 423 GAATCTCTCTGTTTCGGTGACCGTGAATCTATAAGGACTGGTGGAAATGCCAAACTGTT 482
Qy 386 GluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyr 405
Db 483 GAAGAGTACTGGAGGAGTGTGGACATGCTGTTTCATAGTGGATCATCAGACATATAT 542
Qy 406 PheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuVal 425
Db 543 TTTCCATGTATAAGGAAGGCTTTTCCAGGGGTAGCTATTCTAATCTCTGTTCTGTT 602
Qy 426 SerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAla 445
Db 603 TCAGCTGTATTCATGATGATATGATGCGGTGCGGTGCGGCACATTTTCAATTCGTGGCA 662
Qy 446 PheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAnLys 465
Db 663 TTTCTGGGATCATGTTTTCAGATACCGTGGTATCTTGACAGATATCTCCATGCTACG 722
Qy 466 PheArgAnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGln 485
Db 723 TTCAGCATGTATGCTGGCAACATGATATTTGGTTC---TTCAGTATAGTCGGACAG 779
Qy 486 ProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArgLysGlyLys 502
Db 780 CCGATGTGTCTCTTCTATACTACCATGACGTATGACAGGCGGCCAG 830

RESULT 4
BI422326 729 bp mRNA linear EST 16-AUG-2001
LOCUS EST53292 tomato callus, TAMU lycopersicon esculentum cDNA clone
DEFINITION GLEC69N10 5' end, mRNA sequence.
ACCESSION BI422326
VERSION BI422326.1 GI:15196624
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 729)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Hoit,I.E.,
Liang,F., Updon,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute

```

```

Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
            Location/Qualifiers
source     1..729
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="GLEC69N10"
            /clone_lib="tomato callus, TAMU"
            /tissue_type="callus"
            /dev_stage="25-40 days old"
            /lab_host="XLI-Blue MRF"
            /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
            XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
            of seedlings 7-10 days post-germination were excised, cut
            at both ends and placed on MS medium with no selection.
            Mixed callus was harvested at 25 and 40 days and included
            undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 177 a 142 c 160 g 250 t
ORIGIN
Alignment Scores:
Pred. No.:      8.4e-103      Length:      729
Score:          1038.50      Matches:      182
Percent Similarity: 87.85%      Conservative: 35
Best Local Similarity: 73.68%      Mismatches: 26
Query Match:    38.75%      Indels:      5
DB:             13          Gaps:       1
US-09-856-018b-16 (1-504) x BI422326 (1-729)

```

```

Qy 245 ValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSer 264
Db 2 GTGAATGAGGGTGAG-----AATCCGAATCACTACTCTTACATGTTAGT 49
Qy 265 PheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrPro 284
Db 50 TTCGAGAGTTGGCTTACTTTCATGTTGCTCCAACTTTATGCTATCAGCTAGTATCCT 109
Qy 285 ArgThrProTyrIleArgLysGlyTyrPheArgGlnLeuValLysLeuIleIlePhe 304
Db 110 CGCTCTGCATCCATTCGGAAGGGTGGCTGCCGCCCACTCATCAAGCTGGTAATTTT 169
Qy 305 ThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGlnAnSerGln 324
Db 170 ACAGGATTAATGGATTTATCATTTGACGAGTATATTAACCCGATTTGTCGGAAGCTCACA 229
Qy 325 HisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValPro 344
Db 230 CATCCATTTTGAAGGAACCTTTTATAGCCCATCGAGAGGCTATTGAAGCTTTCAAGTTCCA 289
Qy 345 AsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeu 364
Db 290 ATTTTATATGTCGGCTCTGCATGTTCTACAGCTCTTTTCATCTTTGGTTAAATATACTT 349
Qy 365 AlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsnAlaLysThr 384
Db 350 GCGGAAGTTCTCGGATTTGGGATCGTGAGTCTTATAAAGATTTGGTGAACGCAAAAACA 409
Qy 385 ValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeu 404
Db 410 ATTTGATGAGTATTTGGAGACTTTTGGAAATATGCTCTACATAGTGGATGTTTCGTCACATC 469
Qy 405 TyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeu 424
Db 470 TATTTCCATGCTTAAAGAAATGACATACCTAAGGAGTTTGAATGGTGGTCTCTTTCTTT 529
Qy 425 ValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrp 444
Db 530 ATATCTGCTGTTTCCATGAGCTGTGTATTGCTTCTTCTTCTGCGCTATTCAAGTTTGG 589
Qy 445 AlaPheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAnSer 464

```



```

Pred. No.: 1.22e-85 Length: 572
Score: 880.00 Matches: 179
Percent Similarity: 93.40% Conservative: 5
Best Local Similarity: 90.86% Mismatches: 6
Query Match: 32.84% Indels: 7
DB: 13 Gaps: 3

US-09-856-018B-16 (1-504) x BM309699 (1-572)

QY 5 AspGluProGlu-SerValAlaThrAlaLeuAsnHisSerSerLeuArgArgProSe 24
Db 2 GATGAGCCTGATTCGTACCCACTGCTCTCAACCACTCTCCCTGCGCGCGTCCAC 61
QY 24 rAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThrAspSerGlyAs 44
Db 62 CGCC-----GCTGGCCCTCTCAATTCGCCGAGACGACACCCACAGTTCGGTGA 112
QY 44 pAspLeuAlaLysAspSerGlySerAspSerIleAsnSerAspAlaAlaValAs 64
Db 113 TGACTTGGCCAGAGATTCGGTTCGAGACTCCATCAGCAGCGAC-----GCCGCCAA 166
QY 64 nSerGlnGlnAsnGluLysGlnAspThrAspPheSerValLeuLysPheAlaTyAr 84
Db 167 TTCGAACCGCAA---CAAAAACAAGACACTGATTCCTCGCTCTCAAAATTCGCTACCG 223
QY 84 gProSerValProAlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePh 104
Db 224 TCCTTCGCTCCCGCTCATCGCAAGTGGAAGAAAGTCCGCTCAGCTCCGACACCATTTT 283
QY 104 eArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValValAlaValAnSe 124
Db 284 CCCTCAGAGTCACCGCGGCCCTCTCAACCTCTGTATAGTAGTCTTGTCTGTGAATAG 343
QY 124 rArgLeuIleGluAsnLeuMetLysTyGlyTrpLeuIleLysSerGlyPheTrpPh 144
Db 344 CCGACTCATCATTTAATGAAGTATGTTGCTGATCAAAATCTGCTTTGTTGTT 403
QY 144 eSerSerLysSerLeuArgAspPrpProLeuPheMetCysLysSerLeuValValPh 164
Db 404 TAGCTCAAGTCATTCAGAGACTGGCCCTCTTCATGTGTGTCTTCTTCTGTGTATT 463
QY 164 eProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIleProGluProVa 184
Db 464 TCCTTTGTGCTATATATAGTGAGAGAGTGGCAGCAGCAAGTGTATACCCGACACAGN 523
QY 184 lValValValLeuHisIleIleIleThrSerThrSerLeuPheTyPro 200
Db 524 TGTGTGTGGACTTCATATAATATCATACCTCAGCTTCATCTTCTATCCA 572

RESULT 8
BQ998922
LOCUS BQ998922
DEFINITION BQ998922 741 bp mRNA linear EST 22-AUG-2002
            OGG20J23.yg.abl OG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
            OGG20J23, mRNA sequence.
ACCESSION BQ998922
VERSION BQ998922.1 GI:22433318
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE 1 (bases 1 to 741)
AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
            Lin, H., van Damme, M., Lavelle, D., Chevaller, P., Ziegler, J., Ellison
            P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
            Church, S., Jackson, L. and Bradford, K.
            Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compenomics.ucdavis.edu/
            Unpublished (2002)
            Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab

```

```

FEATURES
    source
        1..741
            /location="Qualifiers"
            /organism="Lactuca sativa"
            /cultivar="L.serriola"
            /db_xref="taxon:4236"
            /clone="OGG20J23"
            /lab_host="E.coli"
            /note="Vector: pBRCDNASFIAB: The library was constructed
            from 10 different sources of RNA from a single genotype.
            Separate cDNAs were generated using primers that
            incorporated unique 5' and 3' tags to distinguish each
            source of RNA. cDNAs were then pooled, size-fractionated,
            directionally cloned into a custom medium-copy vector and
            transformations made with four size classes to minimize
            size bias. Details of each source of RNA and library
            construction can be obtained at http://cgpdb.ucdavis.edu/
            TAG_LIB=OG_EFGHJ lettuce serriola
            TAG_TISSUE=flowers pre-fertilized
            TAG_SEQ=GGTTGACGGG"

BASE COUNT 197 a 145 c 144 g 255 t
ORIGIN

Alignment Scores:
Pred. No.: 1.79e-84 Length: 741
Score: 871.00 Matches: 167
Percent Similarity: 81.56% Conservative: 32
Best Local Similarity: 68.44% Mismatches: 45
Query Match: 32.50% Indels: 1
DB: 14 Gaps: 0

US-09-856-018B-16 (1-504) x BQ998922 (1-741)
QY 77 SerValLeuLysPheAlaTyArgProSerValProAlaHisArgLysValLysGluSer 96
Db 10 AGTGTGCACATATGCGTATCGACCGCTCTCTCCAGCTCATCGCAATTAAGAGTCT 69
QY 97 ProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIle 116
Db 70 CCTCTAAGTCTGAGCGCCATTTTCAACGACAGTCAATGAGACTCTTTAACCTTTGCATA 129
QY 117 ValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyGlyTrp 136
Db 130 GTGGTTCATGTGACGCAATGAGAGTCAATGAGACTCATCATCTGAGAATCTGATGAAGTATG 189
QY 137 LeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMet 156
Db 190 TTGATAAATTCCAATTTTGTGTCAGTCAAGATCATGTGAGAGATTCGACACTTCTTATG 249
QY 157 CysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGln 176
Db 250 TGTCGCTCACTCTCTCCAACTTCCCGCTTACTCCCTATATATGTTGAATAATAGCATGG 309
QY 177 ArgLysCysIleProGluProValValValValLeuHisIleIleIleThrSerThrSer 196
Db 310 CAAAAGCATATTTTCAGACCCCTGTTGTAATCACTCTCCATATTCCTAACAACTTACAAC 369
QY 197 LeuPheTrpProValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThr 216
Db 370 ATCTGTATCCAGTTTCATCATCTTTCAGGTTTCAGGTTTCAGGTTTCATCAGGCGTATCA 429
QY 217 LeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyAlaHisThrAsnTy 236
Db 430 TTGATGCTTTGCTGCTGCTAATTAATTTGTTGAGTTGTTGTTGTTGTTGTTGCTGCAATAT 489

```

University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QG_CA_Contig2177, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: QGG20 row: J column: 23.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2003, 17:48:52 ; Search time 311 Seconds
(without alignments)
3649.541 Million cell updates/sec

Title: US-09-856-018b-16
Perfect score: 2680
Sequence: 1 MAISDEPESVATALNHSLR.....OPMCVLLYYHDLNMRKGLKD 504

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09856018/runat_20022003_091205_8317/app_query.fasta_1.647
-DB=N_Geneseq_101002 -OFTM=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -SWAP=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856018 -ECGN_1.1.396 -runat_20022003_091205_8317 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2680	100.0	1942	21	AAA48939	Soybean diacylglyc
2	1821	67.9	1904	21	AAA51482	A. thaliana diacyl
3	1821	67.9	1942	21	AAA88835	Arabidopsis acyl C
4	1821	67.9	1942	21	AAZ45371	Acyl-CoA:cholester
5	1821	67.9	1942	22	AAAS0106	Arabidopsis thalia
6	1803	67.3	1888	21	AAA48932	Arabidopsis diacyl
7	1797.5	67.1	1985	21	AAA51484	A. thaliana AS11 d
8	1701.5	63.5	1975	21	AAA48942	Wheat diacylglycer
9	1622	60.5	1587	21	AAA48938	Rice diacylglycer
10	1328	49.6	5193	21	AAA51483	A. thaliana diacyl
11	1321	49.3	1559	21	AAA48935	Corn diacylglycer
12	1319.5	49.2	5339	21	AAA51485	A. thaliana AS11 d
13	1305	48.7	1281	21	AAA48933	Corn diacylglycer
14	1073.5	40.1	901	21	AAA48936	Corn diacylglycer
15	810.5	30.2	978	21	AAA48934	Corn diacylglycer
16	802.5	29.9	1650	21	AAZ49452	Mouse Diacylglycer
17	800.5	29.9	1976	21	AAA76169	Human ACAT related
18	784.5	29.3	1766	21	AAA88846	Rat acyl CoA:chole
19	784.5	29.3	1766	21	AAZ45385	Acyl-CoA:cholester
20	784.5	29.3	1766	22	AAAS0105	Rat sterol acyltra
21	767.5	28.6	1521	19	AAV01533	Human acylcoenzyme
22	739.5	27.6	1895	21	AAA88842	Human acyl CoA:cho
23	739.5	27.6	1895	21	AAZ45383	DNA encoding a pro
24	614.5	22.9	7490	23	ABL28620	Drosophila melanog
25	610	22.8	380	24	ABQ83322	Arabidopsis thalia
26	585	21.8	993	22	ABA08413	Human cDNA SEQ ID
27	568.5	21.2	629	21	AAA51486	EST with homology
28	568.5	21.2	629	21	AAZ49453	A. thaliana Diacyl
29	545.5	20.4	1122	20	AAAX30335	DNA encoding a hum
30	536	20.0	983	19	AAV01539	Human acylcoenzyme
31	492.5	18.4	470	21	AAA48940	Soybean diacylglyc
32	454.5	17.0	3996	23	ABL28621	Drosophila melanog
33	429.5	16.0	452	24	ABL81866	Human ovarian canc
34	421.5	15.7	774	22	AAAS27244	cDNA encoding nove
35	416.5	15.5	470	24	ABL93444	Arabidopsis thalia
36	408.5	15.2	1885	23	ABL13759	Drosophila melanog
37	397	14.8	3649	19	AAV01536	Human acylcoenzyme
38	396.5	14.8	4260	23	ABL13758	Drosophila melanog
39	396	14.8	4011	15	AAQ63212	Acetyl coenzyme A:
40	393.5	14.7	2657	23	ABL15694	Drosophila melanog
41	390	14.6	2040	21	AAA76170	Human ACAT related
42	389.5	14.5	1509	21	AAZ57360	Human acyl CoA:cho
43	384.5	14.3	542	22	AAZ27552	cDNA encoding nove
44	384	14.3	3650	19	AAAT96368	Human acyl-coenzym
45	376	14.0	1607	21	AAZ57359	Mouse acyl CoA:cho

ALIGNMENTS

RESULT 1
AAA48939
ID AAA48939 standard; cDNA; 1942 BP.
XX
AC AAA48939;
XX
DT 06-DEC-2000 (first entry)
XX
DE Soybean diacylglycerol acyltransferase cDNA #1.
XX
KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
XX triacylglycerol; herbicide; EC2.3.1.20; ss.
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 29..1543
/*tag= a

FT XX /product= Diacylglycerol_acyltransferase

PN WO200032756-A2.

PD 08-JUN-2000.

PF 01-DEC-1999; 99WO-US28354.

XX 02-DEC-1998; 98US-0110602.

PR 31-MAR-1999; 99US-0127111.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB, Kinney AJ, Cahoon RE;

XX WPI; 2000-412308/35.

DR P-PSDB; AAY94519.

XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for
PT synthesis of triacylglycerols and increasing the level of oils in plant
PT seeds

XX Claim 4; Page 51; 62pp; English.

XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,
CC soybean and wheat were screened for sequences with homology to a
CC putative acyl CoA cholesterol acyltransferase related gene from
CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
CC sapiens and Mus musculus. The cDNA clones identified from this process
CC were used to form complete diacylglycerol acyltransferase cDNA
CC sequences. The present sequence is soybean diacylglycerol
CC acyltransferase cDNA from clone srl.pk0098.a8. Diacylglycerol
CC acyltransferases are involved in the synthesis of triacylglycerols.
CC Alteration of the expression of the diacylglycerol acyltransferase
CC cDNA can be useful for increasing the level of oils in plant seeds.
CC Inhibitors of diacylglycerol acyltransferase may be useful as
CC herbicides.

SQ Sequence 1942 BP; 517 A; 425 C; 377 G; 623 T; 0 other;

Alignment Scores:

Pred. No.: 5,17e-275 Length: 1942
Score: 2680.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-856-018b-16 (1-504) x AAA48939 (1-1942)

QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSerLeuArg 20
DB 29 ATGCGGATTTCCGATGAGCTGAAAGTGTAGCCACTGCTCAACCACACTCTCCCTGGC 88
QY 21 ArgArgProSerAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrAsp 40
DB 89 CGCGTCCCTCCGCACTCCACCAGCGGCTCTCAATTGCGCTGAGCAACACCGAC 148
QY 41 SerSerGlyAspAspLeuAlaLysAspSerGlySerAspSerIleAsnSerAspAsp 60
DB 149 AGTTCGGTGATGACTGGCCAGGATTCGTTCCGACGACTCCATCAACAGGAGCAG 208
QY 61 AlaAlaValAsnSerGlnGlnAsnGluLysGlnAspThrAspPheSerValLeuLys 80
DB 209 GCCCGCTCAATTTCCCAACAGCAAAAGCAAAACAGCACTGATTTCTCGCTCTCAAA 268
QY 81 PheAlaThrArgProSerValProAlaHisArgLysValLysGluSerProLeuSerSer 100
DB 269 TTCGCCATCCGCTCCGTCGCCGCTCACGCAAGTGAAGAAAGTCCGCTCAGCTCC 328
QY 101 AspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuVal 120
DB 329 GACACTATTTCCGCTCAGAGTCAGCGGGGCTCTTCAACCTTTGTATAGTAGTCTTGT 388

QY 121 AlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTrpLeuIleLysSer 140
DB 389 GCTGTGAATAGCCGACATCATATTGAGAAATTAATGAAGTATGGTGGTCAATCT 448
QY 141 GlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSer 160
DB 449 GGCCTTTGGTTAGTTCAAGTCATTGAGACATGGCCCTTTTCATGTGTCTTTCT 508
QY 161 LeuValValPheProPheAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180
DB 509 CTTGTGGTATTTCTCTTCGCTGCTTATAGTGAGAAAGTTGGCACAACGGAAGTATA 568
QY 181 ProGluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTrp 200
DB 569 CCCGACACAGTTGTTGTGTACTTATATATACCTCACTCGCTTTCTATCCA 628
QY 201 ValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPhe 220
DB 629 GTTTTAGTTATTCAGGTGTGATTTCTCTTTTGTATCAGGTGTCAAGTTAATGCTGTT 688
QY 221 SerCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAla 240
DB 689 TCTTGTGTGTATGTTTAAATTTGGTCTTATGCACATACAACTATGATGAGAGCA 748
QY 241 LeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrPro 260
DB 749 CTTACCAAAATTAGTTGAAAGGGAGAGCAGCTGCTCGATACTCTGAACATCGACTATCCT 808
QY 261 TyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGln 280
DB 809 TACAACCTAAGCTTCAAGAGCTTGGCATATTTCTGTGTGCTTGCCTACATATTGTTACCA 868
QY 281 ProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLys 300
DB 869 CCAAGCTATCTCCACACCTTATTTGGAAGGGTGGTGTTCGCCAACTTGTCAAG 928
QY 301 LeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleVal 320
DB 929 CTGATAATATTACAGGAGTTATGGGATTTATAATAGACCAATATATTATCCCATAGTA 988
QY 321 GlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLys 340
DB 989 CAAAATTCACAGCATCTCTCAAGGGAAACCTTCTTACGCCACCGAGAGAGTCTTCAAG 1048
QY 341 LeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrp 360
DB 1049 CTTTCTGTTCCTCAAAATTTATATGTGTGCTGTCATGTTCTATTGCTTTTCCACCTTTGG 1108
QY 361 LeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrp 380
DB 1109 TTAATAATCTCTGGCAGAGCTTCTTCGATTTGGTATCGTGAATCTACAAGGATTTGGTG 1168
QY 381 AsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMet 400
DB 1169 AATGCCAAAACCTGCGAAGATTTATGGAGGATGCGAATATGCTGTGTCAAAATGATG 1228
QY 401 IleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeu 420
DB 1229 ATCGCCACCTATATTTTCCATGTTTAAGGCACCGCTCTACCAAGAGCTGCTGCTCTTTA 1288
QY 421 IleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIle 440
DB 1289 ATTGCCCTCTCTGGTTTCTGCTTTATTCATGAGCTGTCATTTGCTTCTTCCATGACATA 1348
QY 441 PheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsn 460
DB 1349 TTCAAGTTGGGGCTTTTCGGTGGAAATTTATGTTTCAGGTTCCTTTGGTCTTGCATCAAT 1408
QY 461 TyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePhe 480
DB 1409 TATCTGCAAAATAAATTCAGAACTCAATGGTGGAAATATGATTTTTTGGTTCATATTC 1468


```
Qy 353 PheTyrCysPhePheHisLeuTyrLeuAsnTleLeuAlaGluLeuLeuLeuArgPheGlyAsp 372
Db 1246 TTCCTACTGCTTCTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTCTGCTCGGGAT 1305
Qy 373 ArgGluPheTyrLysAspTyrPheAsnAlaLysThrValGluAspTyrTyrArgMetTrp 392
Db 1306 CGTGAATTCCTACAAGATGTGTGGAATGCCAAAAGTGTGGAGATTACTGGAGAAATGTGG 1365
Qy 393 AsnMetProValHisLysTyrMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
Db 1366 AATATGCTGTTCATAAATGATGGTTCGACATATATATCTCCGTCGTCGCGAGCAAG 1425
Qy 413 LeuProLysAlaAlaLeuLeuLeuAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1426 ATACCAAGACACTCGCCATTATCATCTTCTAGTCTCTGCAGTCTTTCATGAGCTA 1485
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTyrPalapheGlyGlyIleMetPheGln 452
Db 1486 TGCATCGCAGTTCCTCTCTCTCAAGCTATGGGCTTCTTGGGATTATGTTTCAG 1545
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1546 GTGCTTGTGCTTCATCAACAATACTATCTACAGGAAGTTT---GGCTCAACGTTGGGG 1602
Qy 473 AsnMetIlePheTyrPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
Db 1603 AACATGATCTTCTGTTTCATCTTCTGCTTTCGACAAACCGATGTGTGCTCTTAT 1662
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1663 TACCACGACCTGATGAACCGAAAGGATCGATG 1695
RESULT 3
AAA88835
ID AAA88835 standard; DNA; 1942 BP.
XX
AC AAA88835;
XX
DT 19-FEB-2001 (first entry)
XX
DE Arabidopsis acyl CoA:cholesterol acyltransferase DNA.
XX
KW Acyl CoA:cholesterol acyltransferase; ACAT; transgenic plant;
KW sterol; tocopherol; phytosterol; phytostanol; anticholesterolemic;
KW hypolipemic; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 237..1799
FT FT /*tag= a
FT FT /EC_number= "2.3.1.26"
XX
PN WO200061771-A2.
XX
PD 19-OCT-2000.
XX
PF 12-APR-2000; 2000WO-US09696.
XX
PR 12-APR-1999; 99US-0128995.
XX
PA (MONS ) MONSANTO CO.
XX
PI Venkatramesh M, Corbin DR, Bhat GB, Boddupalli SS, Grebenok RJ;
PI Kishore CM, Lardizabal KD, Lassner MW, Rangvala SH, Karunanandaa B;
XX
DR WPI; 2000-665136/64.
XX
DR P-PSDB; AAB19740.
XX
PT Genetically engineering the biosynthetic pathways in plants involved in
PT the accumulation of sterol compounds and tocopherol to produce
PT compounds for lowering the level of low density lipoprotein cholesterol
```

```
in blood serum -
Disclosure; Page 57-58; 166pp; English.
The present sequence is that of Arabidopsis thaliana DNA encoding
acyl CoA:cholesterol acyltransferase (ACAT, see AAB19740). Sterol
O-acyltransferases such as ACAT catalyze the formation of
cholesterol esters from cholesterol and long chain fatty acids.
Recombinant constructs of the invention are used to alter the
biosynthesis and accumulation of sterols and tocopherols in
transgenic plants. Seeds of such plants may contain elevated
levels of sitostanol and/or its esters, and alpha-tocopherol, and
reduced levels of campesterol and campestanol and their esters.
The seeds may also contain the novel sterol brassicatanol. Oil
obtained from the seeds can be used in food and pharmaceutical
compositions to lower levels of low density lipoprotein cholesterol
in blood serum. ACAT enzymes can be used in the present invention
to produce elevated levels of phytosterol and/or phytostanol esters.
XX
SQ Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;
Alignment Scores:
Pred. No.: 1.6e-183 Length: 1942
Score: 1821.00 Matches: 353
Percent Similarity: 75.33% Conservative: 47
Best Local Similarity: 66.48% Mismatches: 91
Query Match: 67.95% Indels: 40
DB: 21 Gaps: 9
US-09-856-018B-16 (1-504) x AAA88835 (1-1942)
Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSer----- 18
Db 237 ATGGCGATTTTGGATTCT---GCTGGCGTTACTACGGTGGAGACGGTGGCGGAGAG 293
Qy 19 -----LeuArgArgProSer---AlaThrSerThrAlaGlyLeu 31
Db 294 TTCGTCGATCTTGATAGGCTTCGTCGAGGAAATCGAGATCGGATTCTTCTAACGGACTT 353
Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51
Db 354 CTTCTCTCTGTTCCGATAATAATCTCCTTCGGATGATGTTGGAGCTCCCGCCGACGTT 413
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67
Db 414 AGGATCGGATGATTGATTCGGTTGTTAAACGATCAGCTCAGGGAACAGCAATTGGCCGGA 473
Qy 68 GlnAsnGlu-----LysGln 72
Db 474 GATAATAACGGTGTGTCGATAATAACGGTGTGGAAGAGCGCGGAGAAAGAGAGGA 533
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
Db 534 AACCCGATGCTACG-----TTTACGTATCAGCCGCTCGGTTCAGCTCATCGGAGG 584
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112
Db 585 GCGAGAGAGAGTCCACTTAGCTCCGACGCAATCTTCAACAGAGCCATGCCGGATTATTC 644
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132
Db 645 AACCTCTGTGTAGTAGTTCTTATTGCTGTAACAGTAGTACATCATCGAAATCTTATG 704
Qy 133 LysTyrGlyTyrLeuIleLysSerGlyPheThrPheSerLysSerLysSerLeuArgAspTrp 152
Db 705 AAGTATGTTGGTGTATCAGAACGAGATTCTTGGTTTATGTTCAAGATCGCTCCGAGATTGG 764
Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172
Db 765 CCGCTTTTCATGTGTTGTATATATCCCTTTTCGATCTTTCTTTGGCTGCTTTACGGTTGAG 824
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIleIleIle 192
Db 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIleIleIle 192
```

```
Db 825 AATGGTACTTCAGAAATACATATCAGAACCTGTGTGTCATCTTCTTCATATATATATC 884
QY 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
    ||| ||| ::| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 885 ACCATGACAGAGGTTTGTATCCAGTTTACGTTACACCTAAGGTGTGATTCGCTTTTAA 944
QY 213 SerGlyValThrLeuMetLeuPheSerCysValValThrLeuLysLeuValSerTyrAla 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 945 TCAGGTGTCACTTGTATGCTCTCTCACTTCATGTGTGTGCTTAAAGTGTGTTCTTATGCT 1004
QY 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1005 CATACTAGCTATGACATAGATCCCTAGCCAAATGCAGCTGATAAGGCC----- 1052
QY 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272
    ||| ::| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1053 -----AATCCTGAAGTCTCTACTACTAGTTAGCTTGAAGAGCTTGGCATATTTTCATG 1103
QY 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1104 GTGGTCCACATTTGTGTATCAGCAAGTTATCCACGTTCTGCATGTATACGGAAGGTT 1163
QY 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1164 TGGTGGCTCGTCAATTTGCCAAAGCTGGTCATATTCACCGGATTCATGGGATTTAATA 1223
QY 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1224 GAACAATATATAATCTATTGTCAGGAACCTCAAGACATCTTTTGAAGCGCATCTCTA 1283
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1284 TATGCTATTGAAGAGTGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGGCTCTGCATG 1343
QY 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1344 TTCTACTGCTTCTCCACCTTGGTTAAACATATTGGCAGAGCTTCTGCTTCGGGGAT 1403
QY 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1404 CGTAATTCTACAAAGATTTGGTGAATGCAAAAGTGTGGGAGATTACCTGGAGATGTGG 1463
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1464 ATATGCTGTTTCATAAATGGATGGTTCGACATATATATCTCCGCGCTTGCAGCAAG 1523
QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1524 ATACCAAGACACTGCCCATATCATGTTCTTCTAGTCTCTGCAGTCTTTCATGAGCTA 1583
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1584 TGCATCGCAGTTCCTGTGCTCTCTCAAGCTATGGGCTTTCTTGGGATATATGTTTCAG 1643
QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1644 GTGCTTGGTCTTCATCAACAACATCTACAGAAAGGTTT---GGCTCAACGGTGGG 1700
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1701 ACATGATCTCTCTGTTCTATCTCTGCATTTTCGGNACACCGATGTGTGCTCTTTAT 1760
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1761 TACCAGACCTGATGAACCGAAAGGATCGATG 1793
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 4
AAZ45371
ID AAZ45371 standard; DNA; 1942 BP.
XX
XX AAZ45371;
XX
XX 27-MAR-2000 (first entry)
```

```
XX Acyl-CoA:cholesterol acyltransferase (ACAT)-like protein DNA.
DE
XX Acyl-CoA:cholesterol acyltransferase; ACAT; ACAT-like protein; sterol;
KW ester; triacylglycerol; fatty acyl-CoA; lipid composition; plant cell;
KW diacylglycerol acyltransferase; DAGAT; triglyceride; cancer; diabetes;
KW cardiopulmonary disease; heart failure; atherosclerosis; adipocytosis;
KW leukaemia; skin carcinoma; fibroblastoma; metabolic disorder; obesity;
KW abnormal lipid metabolism; abnormal fat absorption;
KW lipoprotein secretion; adipogenesis; ss.
XX
XX Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
FT CDS 237..1799
FT /*tag= a
FT /product= "Acyl-CoA:cholesterol acyltransferase (ACAT)-
FT like protein"
XX
XX WO963096-A2.
XX
XX 09-DEC-1999.
XX
XX 04-JUN-1999; 99WO-US12541.
XX
XX 05-JUN-1998; 98US-0088143.
XX
XX 12-NOV-1998; 98US-0108389.
XX
XX (CALJ ) CALGENE LLC.
XX
XX Lassner MW, Ruzinsky DM;
XX
XX WPI; 2000-105701/09.
XX
XX P-PSDB; AAY54143.
XX
XX Novel polynucleotides used for modifying plant oil composition and for
XX developing products for treating e.g. cancer, diabetes, cardiopulmonary
XX disease or metabolic disorders
XX
XX Claim 4; Fig 1; 89pp; English.
XX
XX The present sequence encodes an acyl-CoA:cholesterol acyltransferase
XX (ACAT) related protein. The ACAT-like protein is active in the formation
XX of a sterol, ester and/or triacylglycerol from a fatty acyl-CoA and
XX sterol and/or diacylglycerol substrate. The DNA can be used for
XX modifying the lipid composition of plant cells. The ACAT-like protein
XX has diacylglycerol acyltransferase (DAGAT) activity, and so the
XX synthesis of triglycerides can be suppressed or increased using the
XX DNA. The protein can be used to produce plant oils with a modified
XX triglyceride content. The products can also be used to identify
XX antagonists and agonists of DAGAT activity. Such agonists and
XX antagonists are particularly useful in treating or ameliorating
XX diseases associated with DAGAT activity, including diseases associated
XX with altered cellular diacylglycerol concentration or PKC activity,
XX atherosclerosis, adipocytosis, leukaemia, skin carcinoma, fibroblastoma,
XX metabolic disorders, obesity, diseases associated with abnormal lipid
XX metabolism, and diseases associated with abnormal fat absorption,
XX lipoprotein secretion and adipogenesis.
XX
XX Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.6e-183 Length: 1942
XX Score: 1821.00 Matches: 353
XX Percent Similarity: 75.33% Conservative: 47
XX Best Local Similarity: 66.48% Mismatches: 91
XX Query Match: 67.95% Indels: 40
XX DB: 21 Gaps: 9
XX
XX US-09-856-018B-16 (1-504) x AAZ45371 (1-1942)
XX
XX 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
```

Db 237 ATGGCGATTTGGATCTCT---GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 293
 Qy 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31
 Db 294 TTCGTCGATCTTGATAGCTTCGTCGACGGAATCGAGATCGGATCTTCTAACGGAGCTT 353
 Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51
 Db 354 CTTCCTCTGCTCGGATATAAATCTCTTCGGATGATGTTGGAGCTCCCGCGACGCTT 413
 Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67
 Db 414 AGGATCGGATGATGATTCGGTTTAAACGATGACGCTCAGGGAACAGCCAATTTGGCCGGA 473
 Qy 68 GlnAsnGlu-----LysGln 72
 Db 474 GATATAACGGTGGTGGCGCATATAAACGGTGGTGAAGAGCGGGGAGAGGAAGAGA 533
 Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
 Db 534 AACCGGATGCTACG-----TTTACGTATCGACCGTGGTTCACGCTCATCGGAG 584
 Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112
 Db 585 GCGAGAGAGTCCACTTAGCTCCGACGCAATCTCAACACGACAGCCATGCCGGATTAATTC 644
 Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132
 Db 645 AACCTCTGTGTAGTAGTCTTATTGCTGTAACAGTAGACTCATCATCGAANAATCTTATG 704
 Qy 133 LysTyrGlyThrLeuIleLysSerGlyPheThrPheSerSerLysSerLeuArgAspTrp 152
 Db 705 AAGTATGTTGGTGTATCAGAACGAGATTTCTGGTTTAGTTTCAAGATCGCTCGGAGATTGG 764
 Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172
 Db 765 CCGCTTTTCATGTGTGTATATCTCTTCGATCTTCTTCGCTGCTTTCACGGTTGAG 824
 Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValLeuHisIleIleIle 192
 Db 825 AAATGGTACTTCAGAAATACATATACAGACCTGTGTCTATCTTCTTCATATATTATC 884
 Qy 193 ThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
 Db 885 ACCATGACAGAGGTTTGTATCCAGTTTACGTCACCTACCCCTAAGGTGTGATCTGCTTTT 944
 Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232
 Db 945 TCAGGTGTCACTTTGATGCTCTCTCACTTGCATTTGTGGCTAAAGTTGGTTTCTTATGCT 1004
 Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
 Db 1005 CATACTAGCTATGACATAAGATCCCTAGCCCAATGCAGCTGATAGGCC----- 1052
 Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272
 Db 1053 -----AATCTCGAGTCTCTACTACCTTAGCTTGAAGAGCTTGGCATATTTTCATG 1103
 Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292
 Db 1104 CTCGCTCCACATGTGTATTACACCAAGATTATCCAGTCTTCGATGTATACGGAAGGGT 1163
 Qy 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312
 Db 1164 TGGGTGGCTCGTCAATTTGCAAAACTGCTCATATTACCGGATTCATGGGATTTATAATA 1223
 Qy 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
 Db 1224 GAACATATATAAATCTTATTGTGAGGAACCTCAAGCAATCTTTGAAAGGCGATCTTCTA 1283
 Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352

Db 1284 TATGCTATTGAAGAGAGTGTTCGAAGCTTTCAGTTCCTCAAAATTTATATGTGTGGCTCTGCATG 1343
 Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
 Db 1344 TTCTACTGCTTCTTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTCTGCTTCGGGAT 1403
 Qy 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
 Db 1404 CGTGAATTTCTACAAGATTGTGGAAATGCAAAAAGTGTGGGAGATTACTGGAGAATGTGG 1463
 Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
 Db 1464 AATATGCTGCTTCAATAATGGATGTTTCACATATATATCTCCGCTGCTTCGCGAGCAAG 1523
 Qy 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
 Db 1524 ATACAAAGACACTCGCCATTTATCTTCTTAGTCTCTGAGTCTTTTCATGAGCTA 1583
 Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
 Db 1584 TGCATCGCAGTTCCTGTGCTCTCTCAAGCTATGGGCTTTTCTTGGGATTTATGTTTCAG 1643
 Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
 Db 1644 GTGCTTTGGTTCATCATCAAACTATCTACAGGAAGGTTT---GGCTCAACGCTGGGG 1700
 Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuTyr 492
 Db 1701 AACATGATCTTCTGGTTCATCTCTGCAATTTTCGACAAACCGATGTGTGCTCTTTAT 1760
 Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
 Db 1761 TACCAGCGACCTGATGAACCGAAAGGATCGATG 1793
 RESULT 5
 AAS01106
 ID AAS01106 standard; cDNA; 1942 BP.
 XX
 AC AAS01106;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Arabidopsis thaliana sterol acyltransferase ACAT EST sequence.
 XX
 KW Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
 KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
 KW nutritional supplement; dairy product; food product; salad dressing;
 KW expressed sequence tag; EST; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 237..1799
 FT /*tag= a
 FT /product= "ACAT"
 XX
 WO200116308-A2.
 PD 08-MAR-2001.
 XX
 PF 30-AUG-2000; 2000WO-US23863.
 XX
 PR 30-AUG-1999; 99US-0152493.
 XX
 PA (MONS) MONSANTO CO.
 XX
 PI Lessner M, Van Eenennaam A;
 DR WPI; 2001-169010/17.
 DR P-PSDB; AAU00462.
 XX
 PT New isolated nucleic acid encoding plant lecithin:cholesterol
 PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol

PT acyltransferase-like polypeptides, for modifying the sterol content and
PT oil production of plants -

PS Claim 4; Page 105-106; 127pp; English.

XX The present sequence encodes for Arabidopsis thaliana
CC acyl CoA:cholesterol acyltransferase-like (ACAT). Several novel
CC polynucleotides encoding the plant sterol acyltransferases LCAT
CC (lecithin:cholesterol acyltransferase-like; AAS001081-AAS01104, AAS01341)
CC and ACAT (AAS01311-AAS01319) are described. A yeast LCAT related open
CC reading frame, LR01 gene sequence (AAS01342), and a rat ACAT (AAS01105)
CC cDNA sequence are also described. The polynucleotides encoding LCAT
CC or ACAT are used to produce LCAT or ACAT polypeptides. They can also
CC be used in a recombinant construct to transform a host cell (preferably
CC of a plant) or a plant. The recombinant construct is used to increase or
CC decrease the sterol content of the host cell or plant. It can be used to
CC alter oil production of the cell or plant, preferably by increasing it.
CC The oil of the plant or the plant itself is used as a food product, or
CC as nutritional or dietary supplements, or in pharmaceutical compositions
CC for lowering cholesterol. The oil can be used in foods e.g. margarine,
CC butter, cooking oil, and dressings e.g. salad dressings, mayonnaise,
CC cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips,
CC baked goods, pastries, cookies, snack bars, confections, chocolates, and
CC beverages. The alteration in sterol content and/or composition can also
CC provide a plant with tolerance to stress and insect damage.

XX SQ Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;

Alignment Scores:

Pred. No.: 1.6e-183 Length: 1942
Score: 1821.00 Matches: 353
Percent Similarity: 75.33% Conservative: 47
Best Local Similarity: 66.48% Mismatches: 91
Query Match: 67.95% Indels: 40
DB: 22 Gaps: 9

US-09-856-018B-16 (1-504) x AAS01106 (1-1942)

QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
DB 237 ATGGCGATTGAGTCTCT---GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 293
QY 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31
DB 294 TTCGTCGATCTGTAGAGGCTTGTGACCGAAATCAGATCGGATCTCTTAACGACAT 353
QY 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspLeuAlaLysAspSerGly 51
DB 354 CTCTCTCTGGTCCGATAAATCTCTCTCGGATGATGTTGGAGCTCCCGCGAGCTT 413
QY 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67
DB 414 AGGGATCGGATGATTCCTGTTTACGATGACGCTCAGGGAACAGCAATTTGGCCGGA 473
QY 68 GlnAsnGlu-----LysGln 72
DB 474 GATAATAACGGTGGCGGATTAATACGGTGGTGGAGAGCGCGGAGAGAGAGA 533
QY 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
DB 534 AACCGCGATGCTACG-----TTTACGTATCGACCGCTCGGTCCACCTCATCGGAG 584
QY 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112
DB 585 GCGAGAGAGAGTCCACTTACGTCGCGCAATCTTCAACACAGAGCCATCGCGGATTATTC 644
QY 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleGluAsnLeuMet 132
DB 645 AACCTCTGTGTAGTCTTATGCTGTAAACAGTAGACTCATCATCGAAAATCTTAG 704
QY 133 LysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152
DB 705 AAGTATGTTGGTTGATCAGACGGATTCTGTTAGTTCAAGATCGCTGCCGAGATTGG 764

QY 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172
DB 765 CCGCTTTTCATGTGTATATCCCTTTCGATCTTCCCTTTCGCTGCCTTTCACGGTTGAG 824
QY 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValValLeuHisIleIle 192
DB 825 AAAATGGTACTTTCAGAAATACATATCAGAACTGTGTGTCATCTTCTTCATATTATTC 884
QY 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
DB 885 ACCATGACAGAGGTTTGTATCCAGTTTACGTCACCCCTAAGCTGTCATCTCTCTTTTA 944
QY 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232
DB 945 TCAGGTGTACCTTGTATGCTCTCTACCTTCATTTGTGTGCTAAAGTTGGTTCTTATGCT 1004
QY 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
DB 1005 CATACTAGCTATGACATAGATCCCTAGCCAAATGACGCTGATAAGGCC----- 1052
QY 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272
DB 1053 -----AATCCTGAAGCTCTCTACTACGTTAGCTTGAAGAGCTTGCATATTTCATG 1103
QY 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292
DB 1104 GTCGCTCCACATGTTGTATCAGCAAGTTATCCACGTTCTGCATGTATACGGAAGGT 1163
QY 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetCysPheIleIle 312
DB 1164 TGGGTGGCTCGTCAATTCGCAAACTGGTTCATATTACCGGATTCATGGGATTTATAATA 1223
QY 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
DB 1224 GAACAATATATAAATCTATTGTCAGGAACCTCAAGACATCTTTGAAAGCGCATCTCTA 1283
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
DB 1284 TATGCTATTGAAGAGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGGCTCTGCATG 1343
QY 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
DB 1344 TTCTACTGCTTCTCCACCTTGGTTAAACATATTTGAGGAGATTTACTGGAGATGTGG 1403
QY 373 ArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
DB 1404 CGTGAATTCACAAAGATTGGTGAATGCAAAAGTGTGGAGATTTACTGGAGATGTGG 1463
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
DB 1464 AATATGCTCTTTCATAAATGATGTTTCGACATATATCTCCCGTCTGCGAGCAAG 1523
QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
DB 1524 ATACAAAGACACTCCCATTTATCTTCTTCTAGTCTCTGACGCTTCTCATGAGCTA 1583
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
DB 1584 TGCATCGCAGTTCCTTGTGCTCTCTTCAAGCTATGGCTTTTCTTGGGATATTATTTTCAG 1643
QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
DB 1644 GTGCGCTTTGGTTTCATCAAACTATCTTACAGGAAGGTTT---GGCTCAACGGTGGG 1700
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
DB 1701 AACATGATCTTCTGTTTCACTCTTCGCAATTTCCGGAACACCGATGTGTGCTCTTTAT 1760
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
DB 1761 TACCAGGACCTGATGAACCGAAGGATCGATG 1793

RESULT 6	AAA48932	AAA48932 standard; cDNA; 1888 BP.	DB	188	TTCTGTCGATCTTGTATAGGCTTCGTCACGGAAATCGAGATCGGATTCCTTCTAACGGACCTT 247
XX	AC	AAA48932;	QY	32	PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51
XX	AC	AAA48932;	DB	248	CTTCTCTCTGGTCCGATATAATTCCTTCGGATGATGTTGGAGCTCCGCGGACGTT 307
XX	DT	06-DEC-2000 (first entry)	QY	52	SerAsp-----AspSerIleAsnSerAspAspAla-----AlaValAsnSerGlnGln 67
XX	DE	Arabidopsis diacylglycerol acyltransferase cDNA.	DB	308	AGGGATCGGATTCGATCCGTTGTTAAACGATGACGCTCAGGGAACAGCAATTTGGCCGGA 367
XX	KW	Diacylglycerol acyltransferase; corn; rice; soybean; wheat;	QY	68	GlnAsnGlu-----LysGln 72
XX	KW	triacylglycerol; herbicide; EC2.3.1.20; ss.	DB	368	GATAATAACGGTGGTGGCGGATAATAACGGTGGTGAAGAGCGCGAGGAAGAGGA 427
XX	OS	Arabidopsis thaliana.	QY	73	AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
XX	FT	Key Location/Qualifiers	DB	428	AACGGCGATGCTACG-----TTTACGTATGACCGCTGGTTCGAGCTCATCGGAGG 478
XX	FT	CDS 131..1693	QY	93	ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 111
XX	FT	/*tag= a	DB	479	GCAGAGAGAGTCCACTTAGCTCCGACGCAATCTTCAACAGAGCCATGCGGATATTTC 538
XX	FT	/product= Diacylglycerol_acyltransferase	QY	113	AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132
XX	OS	WO200032756-A2.	DB	539	AACCTCTGTGTAGTAGTTCTTATTGCTGTAACAGTAGACTCATCATCGAAAAATCTTATG 598
XX	OS	08-JUN-2000.	QY	133	LysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152
XX	PR	01-DEC-1999; 99WO-US28354.	DB	599	AAGTATGGTGGTGTGATCAGACGGATTTCTGGTTAGTTCAAGATCGCTGCAGATGG 658
XX	PR	02-DEC-1998; 98US-0110602.	QY	153	ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172
XX	PR	31-MAR-1999; 99US-0127111.	DB	659	CCGCTTTTCATGTTGGATATCCCTTCGATCTTCCCTTGGCTGCTTACGGTTGAG 718
XX	PA	(DUPO) DU PONT DE NEMOURS & CO E I.	QY	173	LysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIleIleIle 192
XX	PI	Cahoon EB, Kinney AJ, Cahoon RE;	DB	719	AAATGGTATCTTCAGAAATACATATCAGAACCTGTTGGCATCTTCTTCATATATTATTC 778
XX	PI	WPI: 2000-412308/35.	QY	193	ThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
XX	PI	P-PSDB: AAY94512.	DB	779	ACCATGACAGAGTTTGTATCCAGTTTACGTCACCCTAAGGTGTGATTCTGCTTTTATA 838
XX	DR	Polynucleotides encoding diacylglycerol acetyltransferase, useful for	QY	213	SerGlyValThrLeuMetLeuPheSerCysValValThrLeuLysLeuValSerTyrAla 232
XX	DR	synthesis of triacylglycerols and increasing the level of oils in plant	DB	839	TCAGGTGTACATTGATGCTCCCTCCTTCGATTCGATGTTGGTAAAGTTGGTTCTTATGCT 898
XX	PT	seeds -	QY	233	HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
XX	PT	Claim 4; Page 38; 62pp; English.	DB	899	CATACCTAGCTAGCATACATAAGATCCCTAGCAATGCAAGCTGCAAGGCC----- 946
XX	PS	In the present invention, cDNA libraries from Arabidopsis, corn, rice,	QY	253	AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272
XX	PS	soybean and wheat were screened for sequences with homology to a	DB	947	-----AATCCTGAAGTCTCCTACTAGTTAGCTTTGAAGAGCTTGGCATATTTCATG 997
XX	CC	putative acyl CoA cholesterol acyltransferase related gene from	QY	273	ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292
XX	CC	Arabidopsis thaliana and diacylglycerol acyltransferases from Homo	DB	998	GTGCTCCACATGTTGTTATCAGCCCAAGTTATCCAGCTTCTGCATGTATACGAAGGGT 1057
XX	CC	sapiens and Mus musculus. The cDNA clones identified from this process	QY	293	TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312
XX	CC	were used to form complete diacylglycerol acyltransferase cDNA	DB	1058	TGGGTGGCTCGTCAATTTGCAAAACCTGGTTCATATTACCGGATTCATGGGATTTATAATA 1117
XX	CC	sequences. The present sequence is Arabidopsis thaliana diacylglycerol	QY	313	AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
XX	CC	acyltransferase cDNA. Diacylglycerol acyltransferases are involved	DB	1118	GAACAATATATAAATCTATTGTGAGAACTCAAAGCATCTCTTTGAAGGCGATCTCTTA 1177
XX	CC	in the synthesis of triacylglycerols. Alteration of the expression	QY	333	TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
XX	CC	of the diacylglycerol acyltransferase DNA can be useful for	DB	1178	TATGCTATTGAAGAGTGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGGCTCTGCATG 1237
XX	CC	increasing the level of oils in plant seeds. Inhibitors of	QY	353	PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
XX	CC	diacylglycerol acyltransferase may be useful as herbicides.	DB	1238	TTCTACTGCTCTCTCCACCTTTGTTAAACATATATGGCAGAGCTTCTCTCTCTGGGAT 1297
XX	QX	Sequence 1888 BP; 457 A; 394 C; 429 G; 608 T; 0 other;	QY	19	-----LeuArgArgArgProSer-----AlaThrSerThrAlaGlyLeu 31
XX	QX	Alignment Scores:	DB	131	ATGGCGATTTGGATCTCT--GCTGGCGTTTACTAGCGGTGACGGAGACGGTGGCGGAGAG 187
XX	QX	Pred. No.: 1.26e-181 Length: 1888	QY	1	MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
XX	QX	Score: 1803.00 Matches: 351	DB	19	-----LeuArgArgArgProSer-----AlaThrSerThrAlaGlyLeu 31
XX	QX	Percent Similarity: 74.95% Conservative: 47	DB	131	ATGGCGATTT

QY 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
 DB 1298 CBTGAATTCACAAAGATTGGTGAAGTCAAAAGTGGAGATTACCTGGAGATGTGG 1357
 QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
 DB 1358 AATATGCCCTGTTCAATATGGATGGTTCGACATATATATCTCCCGTCTGGCGACGAAG 1417
 QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
 DB 1418 ATACCAAGACACTGCCCATATCATCTTCTTCCTAGTCTCTGCAGTCTTTCATGAGCTA 1477
 QY 433 CysIleAlaValProCysHisIlePheLysLeuTyrAlaPheGlyGlyIleMetPheGln 452
 DB 1478 TGCATCGCAGTTCCTTCTCTCTCAAGCTATGGCTTTCTTGGGATTATGTTTCAG 1537
 QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
 DB 1538 GTGCCCTTTGGTCTTCATCAACAACATCTACAGGAAGGTTT---GGCTCAACGGTGGGG 1594
 QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
 DB 1595 AACATGATCTTCTGGTTCATCTCTGCATTTTCGGACACCGCATGTGTCTCTTTAT 1654
 QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
 DB 1655 TACCAGACCTGATGAACCGAAAGGATCGATG 1687
 RESULT 7
 ID ANA51484 standard; cDNA; 1985 BP.
 AC ANA51484;
 DT 09-OCT-2000 (first entry)
 DE A. thaliana AS11 diacylglycerol acyltransferase cDNA.
 KW DGAT; diacylglycerol acyltransferase; seed oil; fatty acid synthesis;
 KW size; weight; carbon flux; TAG1; insertion mutant; ss.
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 139..1782
 FT /*tag= a
 FT /product= diacylglycerol acyltransferase
 FT /note= "insertion mutant"
 FT misc_RNA 610..690
 FT /*tag= b
 FT /note= "81 bp insertion, duplication of exon 2"
 XX
 PN WO200036114-A1.
 XX
 PD 22-JUN-2000.
 XX
 PF 16-DEC-1999; 99WO-CA01202.
 XX
 PR 17-DEC-1998; 98US-0112812.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Zou J, Taylor DC, Wei Y, Jako CC;
 DR WPI: 2000-431592/37.
 DR P-PSDB; AAY96854.
 XX
 PT New DNA encoding diacylglycerol acyltransferase from Arabidopsis
 PT thaliana for transforming plants and regulating seed oil content, fatty
 PT acid synthesis and seed oil acyl composition in commercial and crop
 PT plants
 XX

PS Claim 5; Page 79; 91pp; English.

XX This cDNA encodes Arabidopsis thaliana ecotype Columbia mutant AS11
 CC diacylglycerol acyltransferase (DGAT). The AS11 mutant TAG1 allele has a
 CC 147 bp insertion located at the central region of intron 2. The insertion
 CC is a duplication of a segment that is composed of 12 bp from the 3' end
 CC of intron 1, the entire sequence of exon 2 (81 bp) and 54 bp from the
 CC 5' end of intron 2. The DGAT and the insertion mutant (AS11) are useful
 CC for regulating seed oil content, the ratio of diacylglycerol to
 CC triacylglycerol proportions in seed oil, fatty acid synthesis, seed oil
 CC acyl composition, seed size/weight and carbon flux into other seed
 CC components in commercial and crop plants. The natural formation of
 CC triacylglycerols can be modified to increase the yield in commercial
 CC plant oils or modify their composition to achieve specific commercial
 CC improvements of plants and plant products.

XX Sequence 1985 BP; 482 A; 415 C; 446 G; 642 T; 0 other;

Alignment Scores:

Pred. No.: 5.25e-181 Length: 1985
 Score: 1797.50 Matches: 353
 Percent Similarity: 71.68% Conservative: 47
 Best Local Similarity: 63.26% Mismatches: 91
 Query Match: 67.07% Indels: 67
 DB: 21 Gaps: 10

US-09-856-018B-16 (1-504) x AAA51484 (1-1985)

QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
 DB 139 ATGGCGATTTTGGATTCT---GCTGGCGTTACTACGGGACGGAACGGTGGCGGAGAG 195
 QY 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31
 DB 196 TTCGTCGATCTTGATAGGCTTCGTCGACGGAATCGAGATCGGATCTTCTCAACGACTT 255
 QY 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51
 DB 256 CTTCTCTCTGGTTCGCGATAATAATCTCTCTCGGATGATGTTGGAGCTCCCGCCGACT 315
 QY 52 SerAsp-----AspSerIleAsnSerAspAspAla-----AlaValAsnSerGlnGln 67
 DB 316 AGGGATCGGATTGATCCGTTGTTACGATGACGCTCAGGGAACAGCCCAATTTGGCCGA 375
 QY 68 GlnAsnGlu-----LysGln 72
 DB 376 GATAATAACGGTGGCGGATAATAACGGTGGTGAAGAGCGCGGAGAGAGAGGA 435
 QY 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
 DB 436 AACGCCGATGCTACG-----TTTACGTATCGACCGCTCGGTTCCAGCTCATCGGAG 486
 QY 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGln----- 106
 DB 487 GCGAGAGAGAGTCCACTTAGCTCCGACGCAATCTTCAACACAGCCATCCCGGATTTATTC 546
 QY 106 ----- 106
 DB 547 AACCTCTGTGTAGTAGTTCTTATTGCTGTAAACAGTAGACTCATCATCGAAATCTTATG 606
 QY 107 ---SerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaValAsnSerArg 125
 DB 607 AAGAGCCCATGCCGATATTATCAACCTCTCTGTAGTAGTTCTTATTGCTGTAAACAGTAG 666
 QY 126 LeuIleIleGluAsnLeuMetLysTyrGlyTyrLeuIleLysSerGlyPheTrpPheSer 145
 DB 667 CTCATCATCGCAAAATCTTATGAAGTATGGTTGGTGTATCGAAGCGGATTTCTGGTTAGT 726
 QY 146 SerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerLeuValValPhePro 165
 DB 727 TCAAGATCGCTCGGAGATGGCGCTTTTCATGCTGTGTATATATCCCTTCGATCTTCCT 786
 QY 166 PheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIleProGluProValVal 185

```
Db 787 TTGGCTGCGCTTTACGGTTGAGAAATGGTACTTCAGAAAATACATATCAGAACCTGTTGTC 846
Qy 186 ValValLeuHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 205
Db 847 ATCTTTCTCATATATATATATATATATATATATATATATATATATATATATATATATATAT 906
Qy 206 ArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSerCysValValTrp 225
Db 907 AGGTGTGATTCGCTTTTATACAGGTGTCTACTTTGATGCTCCTCACTTGCATTGTGTGG 966
Qy 226 LeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuVal 245
Db 967 CTAAGTGTGGTTCTTATGCTCATAGTATGATGATGATGATGATGATGATGATGATGATGAT 1026
Qy 246 GluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSerPhe 265
Db 1027 GATNAGGCC-----ATCTCTGAAGTCCTCTACTAGTTAGTTG 1085
Qy 266 LysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArg 285
Db 1066 AAGAGCTTGGCATATTTTCATGCTCGCTCCACATTTGTTATCAGCCAAGTTATCCACGT 1125
Qy 286 ThrProTyrIleArgLysGlyTyrPheArgGlnLeuValLysLeuLeuIlePheThr 305
Db 1126 TCTGCATGTATACGGAAGGTTGGGTGGCTCGTCAATTTGCAAACTGGTCATATTCACC 1185
Qy 306 GlyValMetGlyPheIleLeuAspGlnTyrIleAsnProIleValGlnAsnSerGlnHis 325
Db 1186 GGATTCATGGATTATATAGACATATATATATATATATATATATATATATATATATATAT 1245
Qy 326 ProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsn 345
Db 1246 CCTTTGAAAGCGCATCTCTATATGCTATTGCAAGAGCTGTGAAGCTTTCAGTTCCAAAT 1305
Qy 346 LeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAla 365
Db 1306 TTATATGTGTGGCTTCGATGTTCTACTGCTTCTCCACCTTTGGTTTAAACATATTCGCA 1365
Qy 366 GluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrVal 385
Db 1366 GAGCTTCTGCTTCGGGATCGGATTTCTACAAAGATTGGTGGATGCAAAAGTTGTG 1425
Qy 386 GluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyr 405
Db 1426 GGAGATTACTGGAGAATCTGGAATATGCCGTTCATAAATGGATGGTTCGACATATATAC 1485
Qy 406 PheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuVal 425
Db 1486 TTCCCGTGTGGCCACAGATACCAAGACACTCGCCATATATCATTTGCTTTCCTAGTC 1545
Qy 426 SerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAla 445
Db 1546 TCTGCAGCTTTTCATGAGCATGATGCATGCGAGTTCCTGTGCTCTCTCAAGCATGGCT 1605
Qy 446 PheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLys 465
Db 1606 TTTTTCGGGATATGTTTTCAGGTGCTTGGCTTTCATCAAACTATCTACAGGAAGA 1665
Qy 466 PheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGln 485
Db 1666 TTT---GGCTCACGGTGGGAACATGATCTTCGGTTCATCTTCTGCATTTTCGGACAA 1722
Qy 486 ProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArgLysLeuLysLeu 503
Db 1723 CCGATGTGTGCTTCTTTATACCAGACCTGATGATGATGATGATGATGATGATGATGATG 1776
RESULT 8
AAA48942
ID AAA48942
AC AAA48942 standard; cDNA; 1975 BP.
XX AAA48942;
XX
```

```
DT 06-DEC-2000 (first entry)
XX
DE Wheat diacylglycerol acyltransferase cDNA #2.
XX
KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
KW triacylglycerol; herbicide; EC2.3.1.20; ss.
XX
OS Triticum aestivum.
OS
FH KEY Location/Qualifiers
FT CDS 107..1633
FT /*tag= a
FT /product= Diacylglycerol_transferase
XX
XX WO200032756-A2.
XX
XX 08-JUN-2000.
XX
XX 01-DEC-1999; 99WO-US28354.
XX
XX 02-DEC-1998; 98US-0110602.
XX
XX 31-MAR-1999; 99US-0127111.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon EB, Kinney AJ, Cahoon RE;
XX
XX WPI: 2000-412308/35.
XX
XX P-PSDB; AAY94522.
XX
XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for
XX synthesis of triacylglycerols and increasing the level of oils in plant
XX seeds
XX
XX Claim 4; Page 56; 62pp; English.
XX
XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,
XX soybean and wheat were screened for sequences with homology to a
XX putative acyl CoA cholesterol acyltransferase related gene from
XX Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
XX sapiens and Mus musculus. The cDNA clones identified from this process
XX were used to form complete diacylglycerol acyltransferase cDNA
XX sequences. The present sequence is wheat diacylglycerol
XX acyltransferase cDNA from clone wrl.pk0119.b6.fis. Diacylglycerol
XX acyltransferases are involved in the synthesis of triacylglycerols.
XX Alteration of the expression of the diacylglycerol acyltransferase
XX cDNA can be useful for increasing the level of oils in plant seeds.
XX Inhibitors of diacylglycerol acyltransferase may be useful as
XX herbicides.
XX
XX SQ Sequence 1975 BP; 463 A; 497 C; 461 G; 550 T; 4 other;
XX
XX Alignment Scores:
XX Pred. No.: 8,75e-171 Length: 1975
XX Score: 1701.50 Matches: 306
XX Percent Similarity: 85.82% Conservative: 57
XX Best Local Similarity: 72.34% Mismatches: 54
XX Query Match: 63.49% Indels: 7
XX DB: 21 Gaps: 2
XX
XX US-09-856-018B-16 (1-504) x AAA48942 (1-1975)
XX
XX Qy 83 TyrArgProSerValProAlaHisArgLysValLysGluSerProLeuSerSerAspThr 102
XX Db 363 TACCGGCGTTCGGCGCGGCC-CACCGCGGGTCAAGGAGAGCCGCTTACTCCGAGCC 421
XX Qy 103 IlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaVal 122
XX Db 422 ATCTTCCGACAGAGCCATGCAGGTCTTCTGAATCTATGATCTTGTGCTGATTGCAGTG 481
XX Qy 123 AsnSerArgLeuIleIleGluAsnLeuMetLysTyrClyTrpLeuIleLysSerGlyPhe 142
XX Db 482 AACAGCAGGCTCATATTCGAGAACTTAAATGAAGTATGGCTTATTAATAGAGCTGGGTTT 541
```

QY 143 TrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerLeuVal 162
Db 542 TGGTTTGTAGTCAAGATCGTGGAGATGGCCACTTCTGATGTGCTCACTTACCC 601
QY 163 ValPheProPheAlaAAlaPheIleValGluLysLeuAlaGlnArgLysCysIleProGlu 182
Db 602 ATTTTCCCACTTGTCTCTCTACAGGAGAGTGGGCTCAAGAAGAGCTCACTCCGTGAT 661
QY 183 ProValValValLeuHisIleIleIleThrSerThrSerLeuPheTyrProValLeu 202
Db 662 CATGTGTCTATTCTTCCATATAATATTACAACCACTGTCTTATCTATCCGGTTGT 721
QY 203 ValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSerCys 222
Db 722 GTGATCTTAAGTGTGAATCAGCAGATATTATCTGGATTGTGTGTTATGTTCAATGCAAGC 781
QY 223 ValValTIPLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThr 242
Db 782 ATTACTTGGTTGAAGCTTGTCTCTTTTGTCTATACAAATATGATATAGGATATTGTCC 841
QY 243 LysLeuValGluLysGlyGlu-----AlaLeuLeuAspThrLeuAsnMetAspTyr 259
Db 842 CAAAGTATTGAAAGGGTCTACACATGCGAGTCTCTATCGATGAGCAAAACATTAAAGGC 901
QY 260 ProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyr 279
Db 902 CCA-----ACTATCAACAGTGTGTGTATTTTCATGTGTGGCCCAACACTTTGTTC 952
QY 280 GlnProSerTyrProArgThrProTyrIleArgLysGlyTyrPhePheArgGlnLeuVal 299
Db 953 CAGCAAGTATCCCGGACAGCATATTATAGAAAGGCTGGGTACCCGCGAGCTTATA 1012
QY 300 LysLeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIle 319
Db 1013 AAATGTGTAGTTTACAGGCTTGATGGGCTTCATAATTCAGCAATACATTAATCAATT 1072
QY 320 ValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeu 339
Db 1073 GTGCAGAAATTCGAAGCATCCATGAACGGAAATTTCTTGATGCTATTGAGAGAGTCTTG 1132
QY 340 LysLeuSerValProAsnLeuTyrValTyrLeuLysMetPheTyrCysPhePheHisLeu 359
Db 1133 AAATCTCAGTCCCAACATATATGATGGCTTGTATGTTCTATTCCTTTTCCATCTG 1192
QY 360 TrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTyr 379
Db 1193 TGGTTCAATATTCTAGCCGAACCTCCCTCTTTTGGTGTGCTGATTCACAAAGGACTGG 1252
QY 380 TrpAsnAlaLysThrValGluAspTyrTyrArgMetTrpAsnMetProValHisLysTyr 399
Db 1253 TGGAAAGCCCAACACAGTTGAAGAGTACTGGGAATGTGGAATATGCTGTTCATAAGTGG 1312
QY 400 MetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeu 419
Db 1313 ATCGTTCGACATATATTTTCCATGCATAGAAGTGGCTTATCAAGGGTGTGCGCAT 1372
QY 420 LeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuLysCysIleAlaValProCysHis 439
Db 1373 CTCATGCCATTTCTGTTTTCAGCTGTATTTCATGAGCTATGATTCGTTCGTCGCCAC 1432
QY 440 IlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThr 459
Db 1433 ATTTTCAAAATATGGGCATTTTCTGGAATCATGTTTTCAGATTCCCTGCTATTCTTGAGC 1492
QY 460 AsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIle 479
Db 1493 AAATATCTTCAAGATAGTTCAGAAATACAAATGGTGGGCAACATGATATTTGGTCTTC 1552
QY 480 PheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArg 499
Db 1553 TTCAGCATATGGGCAACCAATGTGTCTTCTTACTACCATCATCATGATGATGACACAGA 1612

QY 500 LysGlyLys 502
Db 1613 CAGGCTCAG 1621

RESULT 9

AAA48938
ID AAA48938 standard; cDNA; 1587 BP.
XX
AC AAA48938;
XX
DT 06-DEC-2000 (first entry)
XX
DE Rice diacylglycerol acyltransferase cDNA #2.
XX
KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
KW triacylglycerol; herbicide; EC2.3.1.20; ss.
XX
OS Oryza sativa.

Key	Location/Qualifiers
CDS	15..1310
FT	/*tag a
FT	/partial
FT	/note= "The CDS of this sequence only encodes
FT	amino acids 70 to 500 of the protein in AA94518"

WO200032756-A2.

XX 08-JUN-2000.
XX 01-DEC-1999; 99WO-US28354.
XX 02-DEC-1998; 98US-0110602.
XX 31-MAR-1999; 99US-0127111.
XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB, Kinney AJ, Cahoon RE;

XX WPI; 2000-412308/35.

XX P-PSDB; AA94518.

XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for
XX synthesis of triacylglycerols and increasing the level of oils in plant
XX seeds -

XX Claim 4; Page 48-49; 62pp; English.

XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,
XX soybean and wheat were screened for sequences with homology to a
XX putative acyl CoA cholesterol acyltransferase related gene from
XX Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
XX sapiens and Mus musculus. The cDNA clones identified from this process
XX were used to form complete diacylglycerol acyltransferase cDNA
XX sequences. The present sequence is rice diacylglycerol acyltransferase
XX cDNA from clone rls24.pk0034.db.fis. Diacylglycerol acyltransferases are
XX involved in the synthesis of triacylglycerols. Alteration of the
XX expression of the diacylglycerol acyltransferase DNA can be useful for
XX increasing the level of oils in plant seeds. Inhibitors of
XX diacylglycerol acyltransferase may be useful as herbicides.

XX Sequence 1587 BP; 411 A; 323 C; 334 G; 519 T; 0 other;

Alignment Scores:

Pred. No.:	1-85e-162	Length:
Score:	1622.00	Matches: 287
Percent Similarity:	82.12%	Conservative: 62
Best Local Similarity:	67.53%	Mismatches: 74
Query Match:	60.52%	Indels: 2
DB:	21	Gaps: 1

.. US-09-856-018B-16 (1-504) x AAA48938 (1-1587)

Qy	75	AspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLysValLys	94
Db	24	GACTCTCTCCGCG-----TTCAGCTTTCCGCGCGGCGCGGCTGCACCGCAAGCCAAAG	77
Qy	95	GluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeu	114
Db	78	GAGAGCCCCCTCAGCTCCGACGCCATCTTCAGCAGAGAGTCATGCGAGCGCTTTTCAACGCTA	137
Qy	115	CysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyr	134
Db	138	TGCATGTTGTTGTTACTGTCAGTGAACACGAGCTTATATCGAGAACTTAAATCAAGTAT	197
Qy	135	GlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeu	154
Db	198	GGCTATTATAAGAGCTGGGTTTGGTTTAAATGATAAAATCATTCGGGAGCTGGCCACTT	257
Qy	155	PheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValIleLysLeu	174
Db	258	CTAATGTGTTGTTCTAGTCTGCGTCTTCCGCCCTGGGTCATTTGCGAGTTTGAAAGTTG	317
Qy	175	AlaGlnArgLysCysIleProGluProValValValLeuHisIleIleIleThrSer	194
Db	318	GCATTTAAACAATGTTATTACTGATGCTGTGTCTACCTGCCCTCATATCTTCCTTTCAACA	377
Qy	195	ThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheValSerGly	214
Db	378	ACCGAAATTCATATCCAGTGTGTGATCTTAAAGTGTGATCTTCGCAAGTTTGTCTGGC	437
Qy	215	ValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisThr	234
Db	438	TTTTTGTGTATATTATTCGCTGTATGTTTGGCTGGAAGCTGTATCTTTTGCATATACA	497
Qy	235	AsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAsnThr	254
Db	498	AACCATGATATAAGCAACTGACCATCGCGCGCAAGAAGTTGATAATGAACACTAAGCAC	557
Qy	255	LeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAla	274
Db	558	GTTTGACATGGATAATTACAACTCCCAACTTTTATGGGAATCTAATATACTTTCATGGCT	617
Qy	275	ProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeu	294
Db	618	CCATACACTCTGTTATCAGCAAGCTATCCCCGAACCTCATCTGTGTAGAAAAGTTGGCTG	677
Qy	295	PheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAspGln	314
Db	678	ATTTCGTCAAATTTATCTGTACTTGATCTTACTGGCTCTTCAAGGCTTCATTTATTGACAA	737
Qy	315	TyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla	334
Db	738	TACATAAATCCAAATTTGTTGAATTCACAGCATCCATTTGAAAGGAGGACTCCTAAATGCT	797
Qy	335	ThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyr	354
Db	798	GTAGAGACTGTTTGTGAACCTCATACCAATGTTTACCTGTGGCTTTGCATGCTTCAT	857
Qy	355	CysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGlu	374
Db	858	GCATTTTTCACATCTCTGGTTAGATATCTTGCTGAGATCTTTCGATTTTGGTGACCGTAA	917
Qy	375	PheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMet	394
Db	918	TTCTACAAAGATTGGTGAATCCAAAACAATTTGATGAGTATTTGAGAGAAAATGCAATATG	977
Qy	395	ProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuPro	414
Db	978	CCGTGATCAATAATGGGTGCTGCCATATTTACTTTCCTTCGATCGCAAAATGGTATATCA	1037
Qy	415	LysAlaAlaLeuLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIle	434
Db	1038	AAGGAAGTTGCTGCTCTGTATATCATCTCCCTGTTGTTCTCGCGTACTCCAGAGATATGTGC	1097
Qy	435	AlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValPro	454

FT exon /*tag- t 3748..3829
 FT /*tag- u 3830..3933
 FT /*tag- v 3934..3954
 FT /*tag- w 3955..4038
 FT /*tag- x 4039..4101
 FT /*tag- y 4102..4184
 FT /*tag- z 4185..4232
 FT /*tag- aa 4233..4355
 FT /*tag- ab 4356..4418
 FT /*tag- ac 4419..4499
 FT /*tag- ad 4500..4550
 FT /*tag- ae 4551..4675
 FT /*tag- af 4676..4780
 FT /*tag- ag /partial

WO200036114-A1.

22-JUN-2000.

16-DEC-1999; 99WO-CA01202.

17-DEC-1998; 98US-0112812.

(CANA) NAT RES COUNCIL CANADA.

Zou J, Taylor DC, Wei Y, Jako CC;

WPI; 2000-431592/37.

P-PSDB; AA96853.

New DNA encoding diacylglycerol acyltransferase from Arabidopsis thaliana for transforming plants and regulating seed oil content, fatty acid synthesis and seed oil acyl composition in commercial and crop plants

Claim 2; Page 67-68; 91pp; English.

This DNA encodes Arabidopsis thaliana diacylglycerol acyltransferase (DGAT). The DGAT and an insertion mutant (AS11) are useful for regulating seed oil content, the ratio of diacylglycerol/triacylglycerol proportions in seed oil, fatty acid synthesis, seed oil acyl composition, seed size/weight and carbon flux into other seed components in commercial and crop plants. The natural formation of triacylglycerols can be modified to increase the yield in commercial plant oils or modify their composition to achieve specific commercial improvements of plants and plant products.

Sequence 5193 BP; 1582 A; 894 C; 903 G; 1814 T; 0 other;

Alignment Scores:

Pred. No.:	2.26e-130	Length:	5193
Score:	1328.00	Matches:	354
Percent Similarity:	39.78%	Conservative:	49
Best Local Similarity:	34.95%	Mismatches:	96
Query Match:	49.53%	Indels:	517
DB:	21	Gaps:	23

US-09-856-018B-16 (1-504) x AAA51483 (1-5193)

QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18

Db	1761	ATGGCGATTTTGGATTCT---CCTGCGGTTACTACGGTGACGGAACGGTGGCGGAGAG	1817
QY	19	-----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu	31
Db	1818	TTTCGTCGATCTTGATAGGCTTCGTCGACGGAATCGAGATCGATTCTTCTAACGGACTT	1877
QY	32	PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysaspSerGly	51
Db	1878	CTTCTCTCTGGTTCGGATAATAATTCTCTCGGATGATGTGGAGCTCCCGCCGACGTT	1937
QY	52	SerAsp-----AspSerIleAsnSerAspAspAla-----AlaValAsnSerGlnGln	67
Db	1938	ACGGATCGGATTGATTCCTGTTTAAAGATGACGCTCAGGGAACAGCCAATTTGGCCGGA	1997
QY	68	GlnAsnGlu-----LysGln	72
Db	1998	GATAATAACGGTGGTGGCGATAATAACGGTGGTGGAGAGCGCGGAGAGAAGAGGA	2057
QY	73	AspThrAspPheSerValLeuLysPheAlaIleArgProSerValProAlaHisArgLys	92
Db	2058	AAAGCGGATGCTACG-----TTTAGGTATCGACCGCTCGGTTCCAGCTCATCGGAGG	2108
QY	93	ValLysGluSerProLeuSerSerAspThrIlePheArg-----	105
Db	2109	GGGAGAGAGAGTCCACTTAGCTCCGACGCAATCTTCAACACAGGTTTAAATCTCAGAAAT	2168
QY	105	-----	105
Db	2169	CTTCGAATTTGGTGTGTTGCTGTTTATATGGAATGAGTTGGTGGTGTGTTTGA	2228
QY	106	---GlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaValAsnSer	124
Db	2229	TTGCAGAGCCATCGCGGATTATTCAACCTCTGTGTAGTAGTTCTTATTGCTGTAACACT	2288
QY	125	ArgLeuIleIleGluAsnLeuMetLys-----	133
Db	2289	AGACTCATCATCGAAATCTTATGAAGGTTTGTGTTACTTGTCTCTCTTTTAGGAAT	2348
QY	134	-----Ty	134
Db	2349	GAATTCCTTGAATTTATATCAGACGGAATACTTTGTTGTTGCTATCATTCATGTAGTA	2408
QY	134	rGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLe	154
Db	2409	TGGTTGGTTGATCAGAACGGATTCTGTTTGTAGTTCAAGATCGCTCGAGATTGGCCGCT	2468
QY	154	upheMet-----	156
Db	2469	TTTCATGTGTTGGTAAAGAGATGTTTTTATTTCCAGCAATGTATACATTTATACGT	2528
QY	157	-----CysCysLeuSe	160
Db	2529	ATAATGATCAGTTTAGTGATCAAGTTCCTTTGATTCTTCTTTCTTTGTTGGAGTATATC	2588
QY	160	rLeuValValPheProPheAlaPheIleValGluLysLeuAlaGlnArgLysCysI	180
Db	2589	CTTTCGATCTTTCTCTTTGGCTGCCTTTACGGTTGAGAAATGGTACTTCAGAAATACAT	2648
QY	180	eProGluProVal-ValValVal-----	187
Db	2649	ATCAGAACCTGTGATGTAATTAATTTCTCCAGCCATTACTGTAAATTTTATTGAAGACA	2708
QY	188	-----L	188
Db	2709	GTTTGTATCATCAAGAACTTACAAGTTCGTCTTTGAAATGCTCAAGGTTGTCATCTTTC	2768
QY	188	euHisIleIleIleThrSerThrSerLeuPheTrpProValLeuValIleLeu-----	205
Db	2769	TTCATATTATTATCACCATGACAGAGGTTTGTATCCAGTTTACGTCACCCCTAAGGTGAT	2828
QY	205	-----	205

Db	2829	ACTGTTTTCTGGCTCAGCACTTTGTGTGACTGTTTAAAGTTTAAAGTTTAGTTGTGCTGACCCGGTGA	2888
Qy	206	-----ArqCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheS	221
Db	2889	TCITGAAATGGACAGCTGTGATTCGTCTTTTATCAGGTGTACATTTGATGCTCTCTCA	2948
Qy	221	erCysValValTrpLeuLysLeuValSerTyraIaHisThrAsnTyraSpmetArgAlaL	241
Db	2949	CTTGCAATTTGTGGCTAAAGTTGGTTCTTATGCTCATACTAGCTATGACATAAGATCCC	3008
Qy	241	euthrLysLeuValGluLys-----GlyGluAlaLeuLeuAspThr-----	254
Db	3009	TAGCCAATGCAGCTGATAGGTAAATACGAAAAAGACGTATGTAT-AGTCACATTGC	3067
Qy	255	-----LeuAsnMetAspTyrrProt	261
Db	3068	ACTGTGTTACTGTTTTAACCAACACTGTTATGAACATTTAGCCCAATCCCTGAAGTCFCT	3127
Qy	261	yrAsnValSerPheLysSerLeuAlaTyrrPheLeuValAlaProThrLeuCystyr---	279
Db	3128	ACTACGTTAGCTTGAAGAGCTTGGCATATTCATGCTCGCTCCACATTTGTGTTATCAGG	3187
Qy	279	-----	279
Db	3188	TAACTGCAAGTGCACTAACCACTCTTATACTTGCAGAGTTTCTGTCTAAACCTCGGA	3247
Qy	280	-----GlnProSerTyrrProAlqThrProTyrrIleArgLysGlyTrpLeup	295
Db	3248	TCITTGCTTTTCCCAAGCAAGTTATCCAGTTCTGCGATGTATACGAGAGGTGGGTGG	3307
Qy	295	heArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAsp---	313
Db	3308	CTCGTCAATTTGCAAACTGGTCATATTACCGGATTCATGGGATTTAATAGAACAAG	3367
Qy	313	-----	313
Db	3368	TAGCTTTTCACATCTTGCCTTTATTAGTTTTCTTGGTGAAATCATCATCCCTGCGTGT	3427
Qy	314	-----GlnTyrrIleAsnProIleValGln	321
Db	3428	CACCACCTTGACTTCATGTCCTTTTGTACATTTGGCAGTATATAATCTATTGTCAGG	3487
Qy	322	AsnSerGlnHisProLeuLysGlyAsnLeuTyrrAlaThrGluArgValLeuLysLeu	341
Db	3488	AACTCAAGCATCCTTTGAAGCGCATCTCTATATGCTATTGAACAGATGTTGAAGCTT	3547
Qy	342	SerValProAsnLeuTyrrValTrpLeuCysMetPheTyrrCysPhePheHisLeuTrp---	360
Db	3548	TCAGTTCCAAATTTATGTGTGGCTCTGCATGTTCTACTGCTCTTCCACCTTTGGTAT	3607
Qy	360	-----	360
Db	3608	GCTGTGATCCCATCTCTTTCAAAATAAATTTGCCAAATTCGAAAAACCGAAAGCGTAAAT	3667
Qy	360	-----	360
Db	3668	CTCATAGAAATTTGATATTTTTAGTTTCTTAGAGTCGGTGATGTAATTTTCAGTTACTGAA	3727
Qy	361	-----LeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArg	373
Db	3728	CGCAATCTCTTGTCCAAAGGTTAAACATATTTGGCAGAGCTTCTCTGCTTCGGGGATCGT	3787
Qy	374	GluPheTyrrLysAspTrpTrpAsnAlaLysThrValGlu-----	386
Db	3788	GAATTCACAAAGATGTGGTGGAAATGCCAAAGAGTGUGG-AGATGTGACCTATTTTACTCA	3846
Qy	386	-----	386
Db	3847	AAAGAAACTTATGATTTTTTAATGTTGTGCTGTTTTTGGGTCACTTAACCTAACCAAT	3906
Qy	387	-----AspTyrrTrpArgMetTrpAsnMet-----	394
Db	3907	CATGATTACACTGCTCTCTCTTTATCAGTACTGGAGATGTGGAATATGTTATGTTCTCT	3966

Qy	394	-----	-----	394
Db	3967	TCCTAAACATCACCTCTTTGTACACAAAATAGAGAAGACAGCAGCTAAATTAAAGATCTTGT		4026
Qy	395	-----	-----ProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArg	410
Db	4027	TTTCCTTGACAGCCTGTCTAATAATGGATGGTTGCACATATATACTTCCCGTGTTCGCC		4086
Qy	411	HisGlyLeuProLysAlaLa-----	-----	417
Db	4087	AGCAAGATACCAAGGTGAG-TGAGATATATACCGATATACGAATTGTCGAGATTGTGTTTC		4145
Qy	418	-----	-----AlaLeuLeuIleAla	422
Db	4146	TGTGATATAAATTAAACCTCCACACACTTGTGTTTTTCAGACACTCGCCATTTATCATTTGCT		4205
Qy	423	PheLeuValSerAlaLeuPheHisGlu-----	-----	431
Db	4206	TTCTAGTCTCTCGAGCTTTTCATGAGGTATACATACATTTCTACATTCGCCCTGTCCTAG		4265
Qy	431	-----	-----	431
Db	4266	ACGCATCAACACACGCTAGTGAAGAAATGCTAATATTCAAAGCATTGTTTTTACTTTAAAC		4325
Qy	432	-----	-----LeuCysIleAlaValProCysHisIlePhe	441
Db	4326	GATCTGTGTACAAATTTCCCTTTTGACAGCTATGCATCGCAGTTCCCTTGTCTCTCTTC		4385
Qy	442	LysLeuTrpAlaPheGlyIleMetPheGln-----	-----	452
Db	4386	AAGCATGGCGCTTTCTTGGGATATGTTTCAGGTTAAAAAATACTAAACTGCTGCAGT		4445
Qy	453	-----	-----ValPro	454
Db	4446	CGATTTTACTAAACTCTATCTCATATATCTGACCAACCAATTTGTTGAGTAGTGCTCT		4505
Qy	455	LeuValLeuIleThrAsnTyrLeuGlnAsnLysPhe--ArgAsnSerMet-----	-----	470
Db	4506	TTGGTCTTCATCAACAACTATCTACAGGAAGGTTTGGCTCAACGGTATGCTCTCAAAAC		4565
Qy	470	-----	-----	470
Db	4566	CCGAGAAAAATAGAACGAATAACTCTTTCTTCATAGCCTAGCCATTTAAATCGCAATGCT		4625
Qy	471	-----	-----ValGlyAsnM	474
Db	4626	GAACCTTAATAATAAAGGTGATCTGTTTGGAAATGGGATCATATATTATGAGTGGGGAACA		4685
Qy	474	etilePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrH		494
Db	4686	TGATCTTCTGGTTCACTCTCTCGATTTTCGGACACACCGATGTGTGCTCTTTATTATACC		4745
Qy	494	IsAspLeuMetAsnArgLysGlyLysLeu		503
Db	4746	ACGACCTTGATGAACCGAAAAAGGATCGATG		4774

Key	Location/Qualifiers
1	1.1
2	2.1
3	3.1
4	4.1
5	5.1
6	6.1
7	7.1
8	8.1
9	9.1
10	10.1
11	11.1
12	12.1
13	13.1
14	14.1
15	15.1
16	16.1
17	17.1
18	18.1
19	19.1
20	20.1
21	21.1
22	22.1
23	23.1
24	24.1
25	25.1
26	26.1
27	27.1
28	28.1
29	29.1
30	30.1
31	31.1
32	32.1
33	33.1
34	34.1
35	35.1
36	36.1
37	37.1
38	38.1
39	39.1
40	40.1
41	41.1
42	42.1
43	43.1
44	44.1
45	45.1
46	46.1
47	47.1
48	48.1
49	49.1
50	50.1
51	51.1
52	52.1
53	53.1
54	54.1
55	55.1
56	56.1
57	57.1
58	58.1
59	59.1
60	60.1
61	61.1
62	62.1
63	63.1
64	64.1
65	65.1
66	66.1
67	67.1
68	68.1
69	69.1
70	70.1
71	71.1
72	72.1
73	73.1
74	74.1
75	75.1
76	76.1
77	77.1
78	78.1
79	79.1
80	80.1
81	81.1
82	82.1
83	83.1
84	84.1
85	85.1
86	86.1
87	87.1
88	88.1
89	89.1
90	90.1
91	91.1
92	92.1
93	93.1
94	94.1
95	95.1
96	96.1
97	97.1
98	98.1
99	99.1
100	100.1


```

CDS      1..984
FT       /tag= a
FT       /partial
FT       /product= Partial_diacylglycerol_acyltransferase
XX
XX
PN       WO200032756-A2.
XX
XX      08-JUN-2000.
XX
XX      01-DEC-1999; 99WO-US28354.
XX
XX      02-DEC-1998; 98US-0110602.
XX
XX      31-MAR-1999; 99US-0127111.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Cahoon EB, Kinney AJ, Cahoon RE;
XX
XX      WPI; 2000-412308/35.
XX
XX      P-PSDB; AAY94515.
XX
XX      Polynucleotides encoding diacylglycerol acetyltransferase, useful for
XX      synthesis of triacylglycerols and increasing the level of oils in plant
XX      seeds
XX
XX      Claim 4; Page 44-45; 62pp: English.
XX
XX      In the present invention, cDNA libraries from Arabidopsis, corn, rice,
XX      soybean and wheat were screened for sequences with homology to a
XX      putative acyl CoA cholesterol acyltransferase related gene from
XX      Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
XX      sapiens and Mus musculus. The cDNA clones identified from this process
XX      were used to form complete diacylglycerol acyltransferase cDNA
XX      sequences. The present sequence is corn diacylglycerol acyltransferase
XX      cDNA from clone cp1lc.pk005.h23. Diacylglycerol acyltransferases are
XX      involved in the synthesis of triacylglycerols. Alteration of the
XX      expression of the diacylglycerol acyltransferase DNA can be useful for
XX      increasing the level of oils in plant seeds. Inhibitors of
XX      diacylglycerol acyltransferase may be useful as herbicides.
XX
XX      Sequence 1559 BP; 413 A; 299 C; 288 G; 555 T; 4 other;

Alignment Scores:
Pred. No.:      2,06e-130      Length:      1559
Percent:      1321.00      Matches:      235
Similarity:      66.90%      Conservative:      54
Best Local Similarity:      54.40%      Mismatches:      72
Query Match:      49.29%      Indels:      71
DB:      21      Gaps:      1

US-09-856-018B-16 (1-504) x AAA48935 (1-1559)

QY      142 PheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysLysSerLeu 161
      |||||.....: |||||.....: |||||.....: |||||.....: |||||.....: |||||.....:
Db      1 TTTTGGTTTAAATGCTACATCATTCGAGAGATGGCCACTGCTAAATGTGTGCCTTAGTCTA 60

QY      162 ValValPhePropheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIlePro 181
      :||:|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      61 CCCATATTCCCTTGGTGCAATTCGAGTCGAAAGATTGGCATTACAACTCTCGTTAGT 120

QY      182 GluProValValValLeuHisIleIleThrSerThrSerLeuPheTyrProVal 201
      :|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      121 GATCCTGCTACTACCTGTTTTCATCATCTTTTACACATTTTGAATTTGTATATCCAGTG 180

QY      202 LeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSer 221
      |||||.....: |||||.....: |||||.....: |||||.....: |||||.....: |||||.....:
Db      181 CTCGTGATCTTAAAGTGTGATCTTCGCAAGTTTATTCAGCGCTTTGTGTGTGATGTTTATTGCC 240

QY      222 CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeu 241
      ||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      241 TGCATGTTTGGCTGAAGCTTGATGTCATCTTTTGCACATACAAACCATGATATTAAGAAACTG 300

QY      242 ThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyr 261

```


QY	395	-----ProValHisLysTrpMetIleA	402
Db	4147	AGAAAGACGCTAATTAAGATCTGTTTCCCTTGACACGCCCTGTTTCATAAAATGGATGTTCC	4206
QY	402	rgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAla	417
Db	4207	GACATATATACTTCCCGCTGCTTGGCAGCAAGATACCAAGGTGAG-TGAGATATATACC	4265
QY	417	-----	417
Db	4266	GATATGCAATTTGCGAGATTTGTTTCTGTGATATATAATTTAACCCCTCCACACACTTGTGTTT	4325
QY	418	-----AlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGlu	431
Db	4326	TTCAGACACTCGCCATATCATCTTCTTCTAGTCTCTGCGAGTCTTTCATGAGGTATACA	4385
QY	431	-----	431
Db	4386	TACTTCTACATTGCCCTGCTCTAGAGCATGAACACACGCTAGTAGAAAGAAATGCTAA	4445
QY	432	-----LeuC	433
Db	4446	TATTAAGCAATGTTTTTACTTAACGATCTTGCTGTACAAATTTCCCTTTTGACAGCAT	4505
QY	433	ysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln	452
Db	4506	GCATCGCAGTTCCTTCTGCTCTTCAAGCTATGGCTTTCTTGGGATATGTTTCAGG	4565
QY	452	-----	452
Db	4566	TTAAAAAATTACTAAACTGCTGCAGTCGATTTTTACTAAACTCTAATCTCATATTTCTGAC	4625
QY	453	-----ValProLeuLeuIleThrAsnTyrIleuGlnAsnLysP	466
Db	4626	CAACCAATTTGTTGATGAGTGGCTTTGGTCTTCATCACAAACTATCTACAGGAAGGT	4685
QY	466	he--ArgAsnSerMet	470
Db	4686	TTGGCTCAACGGTATGCTCTCAAAACCCGAGAAAAATAGAACGAATAACTCTTCTTTTCAT	4745
QY	470	-----	470
Db	4746	AGCTAGCCATTTAAATCCGATGCTGAACCTTAATAATAAAGGTGATCTGTTTGGAAAT	4805
QY	471	-----ValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGln	485
Db	4806	GGGATCATATATTAGTGGGGAACATGATCTCTGGTTCATCTTCGATCTTCGATCTTCGACA	4865
QY	485	nProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArgLysGlyLysLeu	503
Db	4866	ACCGATGTGTGCTGCTTTTATTACACGACCTGATGAACCGAAAGGATCGCATG	4920
RESULT 13			
AAA48933			
ID	AAA48933 standard; cDNA; 1281 BP.		
XX	AAA48933;		
AC	(first entry)		
DT	06-DEC-2000		
XX	Corn diacylglycerol acyltransferase cDNA #1.		
DE	Diacylglycerol acyltransferase; corn; rice; soybean; wheat;		
XX	triacylglycerol; herbicide; EC2.3.1.20; ss.		
KW	Zea mays.		
XX	Location/Qualifiers		
OS	1..1090		
XX	/tag= a		
PH	/partial		
FT	/transl_except= (pos:283..286,aa:ile)		
FT			

[illegible]

Db 241 GTTGGCTGAGCTGTATCTTTTGCACATACAAACCATGATATATAAGGAAACTGATCAC 300
 QY 244 LeuValGluLyGly-GluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnVa 263
 Db 301 AAGCGCAAGAAGGTTGATAATGAATGACCGCGGTGCATAGATAATTTACAANCTCC 360
 QY 263 lSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGln-Pro-Ser 282
 Db 361 AACTCTTGGGAGTCAACATCTCAAGATGGCTCCGACACTCTGTATCAAGCCAAAGT 420
 QY 283 TyrPro-ArgThrProTyrIleArgLysGlyTyrPheLeuAspThrLeuValLysLeu1 302
 Db 421 TATCTNCCGACACCTTATGTAGAAAGTTGGTGGTCCGTCAGATTATTTCTATCTT 480
 QY 302 ellePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGlnAs 322
 Db 481 GATATTTACTGCTCTCCAAGGATTCTATTATGAGCAATACATAAATCCCTATTGTGAA 540
 QY 322 nSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSe 342
 Db 541 CTCTCAACATCCATTGATGGGAGATTACTGAATGCTGTAGAGACTGTTTTGAAGCTCTC 600
 QY 342 rValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAs 362
 Db 601 ATTACCAATGCTACCTGGCTTTCATGTTTATTGCTTTCCTATCTGGTTAA 660
 QY 362 nIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsnAl 382
 Db 661 CATACTTGTGAGATTCTTCGATTTGGTGACGAGAAATCTACAAGACTGGTGGAAATGC 720
 QY 382 aLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIleAr 402
 Db 721 AAGACATGTAGTACTGAGAAATGGAACATGCTGTGCATAAATGATTGTTCCG 780
 QY 402 gHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaIleAlaLeuLeuIleAl 422
 Db 781 TCATATATATTTCCCTTGCATCGGAATGATATCAAGGAAGTTGCTGTTTTATATC 840
 QY 422 aPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIlePheLy 442
 Db 841 GTTCTTTGTTTCTGTACTTACTGATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
 QY 442 sLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLe 462
 Db 901 GTTCTGGGCTTCTTAGGAATCATGCTTCAGATTCCTTCATCATATGACATCATACCT 960
 QY 462 uGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerI1 482
 Db 961 CAAAATAAATTCAGTGACACAAATGTTGGCAATATGATCTTTTGGTTTTTCTGCAT 1020
 QY 482 eLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArgLysGlyLy 502
 Db 1021 ATACGGGCGCAATGTGTCTATTGTATTGATTACCATGATGATGATGATGATGATGATG 1080
 QY 502 s 502
 Db 1081 G 1081
 RESULT 14
 AAA48936
 ID AAA48936 standard; cDNA; 901 BP.
 XX
 AC AAA48936;
 XX
 DT 06-DEC-2000 (first entry)
 DE
 DE Corn diacylglycerol acyltransferase cDNA #4.
 XX
 KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
 XX triacylglycerol; herbicide; EC2.3.1.20; ss.
 OS Zea mays.
 XX

FH Key Location/Qualifiers
 FT CDS 1..858
 FT /*tag= a
 FT /partial
 FT /transl_except= (pos:442..444,aa:Xaa)
 FT /transl_except= (pos:490..492,aa:Xaa)
 FT /transl_except= (pos:587..589,aa:Xaa)
 FT /transl_except= (pos:805..807,aa:Xaa)
 FT /transl_except= (pos:820..822,aa:Xaa)
 FT /note= "Xaa= unknown"
 FT /product= Partial_diacylglycerol_acyltransferase
 XX
 PN WO200032756-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 01-DEC-1999; 99WO-US28354.
 XX
 PR 02-DEC-1998; 98US-0110602.
 PR 31-MAR-1999; 99US-0127111.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Cahoon EB, Kinney AJ, Cahoon RE;
 XX
 DR WPI; 2000-412308/35.
 DR P-PSDB; AAY94516.
 XX
 PT Polynucleotides encoding diacylglycerol acetyltransferase, useful for
 PT synthesis of triacylglycerols and increasing the level of oils in plant
 PT seeds -
 XX
 PS Claim 18; Page 46; 62pp; English.
 XX
 CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,
 CC soybean and wheat were screened for sequences with homology to a
 CC putative acyl CoA cholesterol acyltransferase related gene from
 CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
 CC sapiens and Mus musculus. The cDNA clones identified from this process
 CC were used to form complete diacylglycerol acyltransferase cDNA
 CC sequences. The present sequence is corn diacylglycerol acyltransferase
 CC cDNA from a contig of clones p0042, cspaf49, p0122 ckamb57r and
 CC p0125.czaau61rb. Diacylglycerol acyltransferases are involved in the
 CC synthesis of triacylglycerols. Alteration of the expression of the
 CC diacylglycerol acyltransferase DNA can be useful for increasing the
 CC level of oils in plant seeds. Inhibitors of diacylglycerol
 CC acyltransferase may be useful as herbicides.
 XX
 SQ Sequence 901 BP; 256 A; 172 C; 193 G; 276 T; 4 other;
 Alignment Scores:
 Pred. No.: 2,07e-104 Length: 901
 Score: 1073.50 Matches: 194
 Percent Similarity: 82.73% Conservative: 36
 Best Local Similarity: 69.78% Mismatches: 40
 Query Match: 40.08% Indels: 8
 DB: 21 Gaps: 3
 US-09-856-018B-16 (1-504) x AAA48936 (1-901)
 QY 229 ValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGly 248
 Db 31 GTCTTTATGCACATACAAATATGATATGAAGGGTATTGTCACAAAGTACTGAGAAAGGT 90
 QY 249 GluAla-----LeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSerPhe 265
 Db 91 GCTGCATATGGAATATGTCGATCTCGATATGAAAGATCCA-----ACCTTT 141
 QY 266 LysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArg 285
 Db 142 AAAAGTCTAGTGTACTTCTATGTTGGCCCCAACACTTTGTTACGAGCAACTTATCTCA 201
 QY 286 ThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIlePheThr 305


```
Db 489 CCAGTCCCTTCGAGAAATTATGATACATTTTGTTCGAATTGTACACCAGTTATGTGT 548
Qy 434 ealaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValPr 454
Db :|||||
Db 549 TGCCTGTCCTGCCACATACTCAAGTCTCGGCTTCTTAGGAATCATGCTTCAGATTCC 608
Qy 454 oLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMe 474
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 609 CCTCATATATTGACATCATACCTCAAAATAAATTCACTGACACAATG----- 657
Qy 474 tilePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeuTyrTyrHI 494
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 658 -----CCAATGTGTGTCTTCTATTGTATTACCA 683
Qy 494 sAspLeuMetAsnArgLysGlyLys 502
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 684 TGATGTGATGAACCGGACTGAGAAG 708
```

Search completed: February 21, 2003, 19:51:58
Job time : 348 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 21, 2003, 18:31:35 ; Search time 3369 Seconds
(without alignments)
4353.759 Million cell updates/sec

Title: US-09-856-018b-16
Perfect score: 2680
Sequence: 1 MAISDEPESVATALNHSSLR.....QPMCVLLYYHDLNMRKGKLD 504

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlpl
-Q=/cgn2_1/USPTO_spool/US09856018/runat_20022003_091205_8339/app_query.fasta_1.647
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856018@cgn_1_1_3745-brnat_20022003_091205_8339 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.or.*
21: em.ov.*
22: em.pat.*
23: em.ph.*
24: em.pl.*
25: em.ro.*
26: em.sts.*
27: em.un.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rnd.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1943	72.5	1964	8	AF298815	AF298815 Perilla f
2	1871	69.8	2099	8	AF129003	AF129003 Nicotiana
3	1861	69.4	1537	8	AF251794	AF251794 Brassica
4	1836	68.5	2090	8	AY084052	AY084052 Tropaeolu
5	1821	67.9	1845	8	AF051849	AF051849 Arabidops
6	1821	67.9	1904	8	ATH238008	ATH238008 Arabidops
7	1821	67.9	1942	6	AX090345	AX090345 Sequence
8	1821	67.9	1942	6	AX090349	AX090349 Sequence
9	1821	67.9	1988	8	ATH131831	ATH131831 Arabidops
10	1821	67.9	2005	8	AY054480	AY054480 Arabidops
11	1794.5	67.0	1512	8	AF164434	AF164434 Brassica
12	1519	56.7	1446	8	AF155224	AF155224 Brassica
13	1328	49.6	92822	8	AC005317	AC005317 Arabidops
14	1328	49.6	97495	8	AC003058	AC003058 Arabidops
15	807.5	30.1	1650	10	AF078752	AF078752 Mus muscu
16	807.5	30.1	1776	10	BC003717	BC003717 Mus muscu
17	800.5	29.9	1976	9	AF059202	AF059202 Homo sapi
18	800.5	29.9	1998	9	BC015762	BC015762 Homo sapi
19	795.5	29.7	1792	9	AF236018	AF236018 Cercopith
20	784.5	29.3	1766	6	AX090340	AX090340 Sequence
21	779.5	29.1	1751	10	AF296131	AF296131 Rattus no
22	769.5	28.7	1935	4	AY093657	AY093657 Sus scrof
23	764.5	28.5	1493	9	BC006263	BC006263 Homo sapi
24	753	28.1	185932	2	AP003714	AP003714 Oryza sat
25	752	28.1	1497	3	AF221132	AF221132 Caenorhab
26	739.5	27.6	1895	6	AX090339	AX090339 Sequence
27	739.5	27.6	2074	3	AY051835	AY051835 Drosophil
28	734.5	27.4	2117	3	AF468650	AF468650 Drosophil
29	734.5	27.4	2342	3	AF468649	AF468649 Drosophil
30	678.5	25.3	34948	3	CBRG33E23	CBRG33E23
31	632	23.6	11966	3	CEH19N07	CEH19N07
32	614.5	22.9	40494	2	AC005122	AC005122 Drosophil
33	614.5	22.9	60144	2	AC019798	AC019798 Drosophil
34	614.5	22.9	91685	2	AC005121	AC005121 Drosophil
35	614.5	22.9	170362	3	AC009250	AC009250 Drosophil
36	614.5	22.9	258166	3	AE003652	AE003652 Drosophil
37	570	21.3	728	10	RNO345014	RNO345014 Rattus no
38	546.5	20.4	253305	3	PFMAL3P7	PFMAL3P7
39	420	15.7	262797	3	CNS07EGG	CNS07EGG
40	410	15.3	8593	4	AY065621	AY065621 Bos tauru
41	408.5	15.2	2102	3	AY113537	AY113537 Drosophil
42	406	15.1	1750	10	D86373	D86373 Rattus norv
43	405	15.1	645	11	PMIC2B	PMIC2B
44	400	14.9	3041	10	S81092	S81092 acyl-coenzy
45	400	14.9	3697	10	MUSACACT	L42293 Mus musculu

ALIGNMENTS

RESULT 1

AF298815 1964 bp mRNA linear PLN 16-OCT-2000
 LOCUS Perilla frutescens diacylglycerol acyltransferase (dgatl) mRNA,
 DEFINITION complete cds.
 ACCESSION AF298815
 VERSION AF298815.1 GI:10803052
 KEYWORDS
 SOURCE Perilla frutescens.
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
 1 (bases 1 to 1964)
 Hwang,S.-K. and Hwang,Y.-S.
 Isolation of Perilla frutescens diacylglycerol acyltransferase cDNA
 Unpublished
 2 (bases 1 to 1964)
 Hwang,S.-K. and Hwang,Y.-S.
 Direct Submission
 Submitted (23-AUG-2000) Division of Biochemistry, NIAST, RDA, 249
 Seodun-dong, Gwonsun-gu, Suwon 441-707, Korea
 FEATURES
 source
 1..1964
 /organism="Perilla frutescens"
 /cultivar="Okdong"
 /db_xref="taxon:48386"
 1..1964
 /gene="Dgat1"
 69..1673
 /gene="Dgat1"
 /note="Dgat"
 /codon_start=1
 /product="diacylglycerol acyltransferase"
 /protein_id="AAG23696.1"
 /db_xref="GI:10803053"
 /translation="MAILDSPEILDTSSADNGAAHTTLRRQSRVPPPLDSDS
 NSLEASINDENRDNLIENRGAVSENEKQESYKKEEGAKVNGENSTNGN
 GTDMVAKFTFRAAPAHNRKNKSPSSDAIFKOSHAGLFNLCIVLVAYNSRLIEN
 LMYGNLIGKSFWSSTLSRDLPLMLCCLSLPFAFASFLVLEKLVKLYIPWVAVFL
 HVTITVTEILFPVVDYILRCDASVLSVTLVWLVKLVSAHYAHNYDLRVLAKSID
 KWEMSRMYLDAYDVYDFKSLAYFVAVPTLCYQPSYPTACIRKGVVRLIKLVIF
 TLMGFTIEQYINPIVONSQHLNLLYALERVLLKLSVNLVYWLCMFCFPHLNLN
 ILAELLCFGDREYKQDNWNRATVEEYRWMMNPVHKWVRHITCPCLONGIPKIVAVL
 IALVSAIFHELCAVPCQIFKFWAFSGIMLQVPLVIVITVTLQEKFKNSVMVGNMFWC
 FFCFGQPMCVLLYHYHDLNMRKASAR"
 BASE COUNT 522 a 395 c 461 g 586 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..51e-162 Length: 1964
 Score: 1943.00 Matches: 369
 Percent Similarity: 77.74% Conservative: 50
 Best Local Similarity: 68.46% Mismatches: 72
 Query Match: 72.50% Indels: 48
 DB: 8 Gaps: 6
 US-09-856-018b-16 (1-504) x AF298815 (1-1964)
 Qy 1 MetAlaIleSerAspGluProGluSerValAlaThr----- 12
 Db 69 ATGGCGATCTGGACTCCCGGAGATCTCGACACGAGCTGCTCCAGTGGCGGACACGGC 128
 Qy 13 AlaLeuAsnHisSerSerLeuArgArgProSerAlaThrSerThrAlaGlyLeuPhe 32
 Db 129 GCCGCACATCACACCACTCTTCGCGGAGACAAAGTGGCGCTCCGTTCCGCCCTTCTC 188
 Qy 33 AsnSerProGluThrThrAspSerSerGlyAspPheLeuAlaLysAspSerGlySer 52
 Db 189 -----GATCCGATTCCAACTCTCTGTGAGGAGAGAGCGCAATC 227
 Qy 53 AspAspSerIleAsn-----SerAspAspAla-----Ala 62
 Db 228 AATGATCCGGAATATGTCGAACAGACGCTAATTTGATCGAATATCTCCGCGGCGGAGCC 287

Qy 63 ValAsnSerGlnGlnGlnAsnGlnLysGlnAsp----- 73
 Db 288 GTGGAATCC-----GAGAACGAAAAACAGAGAGCTTATGTTAAGGAGAGGGGCGGAAA 341
 Qy 74 -----ThrAspPheSerValLeuLysPhe 81
 Db 342 GTGAAGGAGAAATGGAGAAACTAGTAATGGCAACGGAACCTGATGTTATGGCGCTCAAAATTC 401
 Qy 82 AlaTyrArgProSerValProAlaHisArgLysValLysGluSerProLeuSerSerAsp 101
 Db 402 ACATTTCAGCGCGCGCGCGCTGCTCACGCAAAAATAAGGAGAGTCTCTTAGCTCCGAC 461
 Qy 102 ThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAla 121
 Db 462 CCATCTTCAACACAGACGCATCGAGGCCCTCTTCAACCTTTGTATAGTGGTCTGTGCT 521
 Qy 122 ValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTyrLeuIleLysSerGly 141
 Db 522 GTAAATAGCAGACTAATAATTTAGAAATTTAATGAAGTATGGTGGCTGATCAAAATCAGGA 581
 Qy 142 PheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysLysLeuSerLeu 161
 Db 582 TTTTGGTTAGTTCAACATCGCTTAGGGAATTGGCCACTGCTTAATGTGTGTCTTAGTCTT 641
 Qy 162 ValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIlePro 181
 Db 642 CCAGTTTTCACCTCGCTTCATTTCTTGTGCGAAGATTGGTGAACACTAAATATATACCT 701
 Qy 182 GluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTyrProVal 201
 Db 702 GAGTGGGTGCGCAGTCTTCTTCTCATGTTACAATCACAACAGTGGAAATCTTGTCTCCAGTT 761
 Qy 202 LeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSer 221
 Db 762 GTTGTCTATCTTAGTGTGATCTGCTGTTCTATCAGGTGTACGCTTAATCTCTTGTCT 821
 Qy 222 CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeu 241
 Db 822 TGCACCTGTATGGTTGAAGCTGTTTCTTCCACGCACATACAAACTATGATTTGAGAGTACTT 881
 Qy 242 ThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyr 261
 Db 882 GCAAAATACACTTGATAGTGGGAAGTACTGCCAGGTACTGGAACCTCGACTACGCTTAT 941
 Qy 262 AsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnPro 281
 Db 942 GATGTAACTTTAAGAGTCTGGCATACTTCATGTTGCTCTCATATTGTGTACAGGCA 1001
 Qy 282 SerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeu 301
 Db 1002 AGCTACCTCGGACAGCTTGCATTTCGGAAGGTTGGTGGTAAAGCAACTAAATTAAGCTG 1061
 Qy 302 IleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGln 321
 Db 1062 GTAAATATTCAGAGACTCATGGGATTTATTATAGAACAGTACATAAAACCCGATCGTTCAA 1121
 Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeu 341
 Db 1122 AATTCTCAACATCTCTGAAAGGAAACCTTTTATGCCATTGAGAGGGCTTGAAGCTT 1181
 Qy 342 SerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeu 361
 Db 1182 TCTCTTCCAAATTTATATGTGTGGCTCTGCATGTTTATTTGTTTTTCCACCTCTGGCTA 1241
 Qy 362 AsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsn 381
 Db 1242 AATATCTTCTGTAACCTTCTGTCTTTGGGACCGTGGATTTTATAAGGATTTGGTGAAT 1301
 Qy 382 AlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTyrMetIle 401
 Db 1302 GCGAGGACAGTGGAGGAGTACTGGAGAAATGTGGAAATATGCCTGTCCATAAATGATGGTT 1361


```
Qy 269 AlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyr 288
Db 1096 GCTTACTTCATGGTTGGCCCAACATTTATGTTATCAGCTTATCCACACATCCATGC 1155
Qy 289 IleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIlePheThrGlyValMet 308
Db 1156 ATTCCGAAGGTTGGTGGCCAGCAATTCATCAAGCTGGTAATATTACAGGATTGATG 1215
Qy 309 GlyPheIleLeuAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLys 328
Db 1216 GGATTATCATAGAACAGTACATTAACCAATGTGCNAACACTCACACATCCCTTTGAAA 1275
Qy 329 GlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrVal 348
Db 1276 GGAACCTTTTATATGCCATCGACAGGGTATTGAAGCTTTTCGGTTCCAAATTTATATGTC 1335
Qy 349 TrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeu 368
Db 1336 TGGCTCTGCATGTTTACTGCTTCTTTCATCTTTGGCTAATATATCTTGCAGAACTACTA 1395
Qy 369 ArgPheGlyAspArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyr 388
Db 1396 TGTTTGGTGATCGTGAGTCTACAGGATGGTGGATGCGCAACAAATGTAGATGATAC 1455
Qy 389 TrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCys 408
Db 1456 TGGAGGATGTGGAATATGCTGTTTCATAGTGGATGTTTCGTCACATTTATTTCCCTTGC 1515
Qy 409 LeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeu 428
Db 1516 TTAAGAACGGAATTCAAAGGGGTGCAATCTGATGATGCTTTCCTTGTATCTGCTGTT 1575
Qy 429 PheHisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGly 448
Db 1576 TTCCACGAGCTGTATGCTGTTCCATGTCGCTTTCAGTGGTGGCATTCATGGGA 1635
Qy 449 IleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsn 468
Db 1636 ATTATGTTCCAGGTTCTTGTGTCATATCACAACATCTTACAAAACAGTTCCAAAGC 1695
Qy 469 SerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlnProMetCys 488
Db 1696 TCGATGTTGGCAATATGATGTTCTGCTGCTTTTCTGCAATCTTGTGTCAGCAATGTTGT 1755
Qy 489 ValLeuLeuTyrTyrHisAspLeuMetAsnArgLys 500
Db 1756 GTGCTTCTGTATTACCAGCATGTGATGAATAGAAAA 1791
```

RESULT 3

```
AF251794 LOCUS 1537 bp mRNA linear PLN 16-APR-2000
DEFINITION Brassica napus putative diacylglycerol acyltransferase mRNA,
complete cds.
ACCESSION AF251794
VERSION AF251794.1 GI:7576940
KEYWORDS
SOURCE Brassica napus.
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 1537)
REFERENCE Brown,A.P., Schierer,T.P. and Slabas,A.R.
AUTHORS Characterization of a putative diacylglycerol acyltransferase cDNA
TITLE from Brassica napus embryo
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1537)
AUTHORS Brown,A.P., Schierer,T.P. and Slabas,A.R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2000) Biological Sciences, University of Durham,
South Road, Durham DH1 3LE, UK
FEATURES
source Location/Qualifiers
1..1537
```

```
CDs
/organism="Brassica napus"
/cultivar="Commanche"
/db_xref="taxon:3708"
/dev_stage="embryo"
22..1527
/EC_number="2.3.1.20"
/codon_start=1
/product="putative diacylglycerol acyltransferase"
/protein_id="AA564065.1"
/db_xref="GI:7576941"
/translational="MEILDSGGVTMPTEGGADLTLRHKPRSDSSNGLLPDSVTVS
DADVRDVSADVEDTQKLANLAGEIRESGGAGGNDVRYTYRPSVPAHRRVRES
LSSDAIFKQSHAGLNLVAVNSRLIENLMKYGLIRTDWFSSSLRDWPLF
MCCLSLRPLAAFTVEKLVOKCISEPVVILHILITTEVLYPVVTLRCDAPLS
GVTMLLTCTVWLKLYSAHTNYDIRTANSKDKANPEVSYVLSLAYPLAPLIC
YQSPSPSCIRKQWARGAKLIIFTGMFIIETQINPVRNSKHPKAGDLLYGYE
RVKLKSVPNVYVWLCMFYCFHLMNLIAELLCGDFREYKDKMNAKSDYWRMNM
VHKWVRHVYFPLRRNIPKPAIIAFLAVSAVHELCIAVPCRLFKWAFLGIMFO
VPLVFTNTYLQERFGSMVNMIEFWFFCFIQGPQMCVLLYYHDLNMRKGMKS"
BASE COUNT 371 a 322 c 355 g 489 t
ORIGIN
```

Alignment Scores:

```
Pred. No.: 2,05e-155 Length: 1537
Score: 1861.00 Matches: 344
Percent Similarity: 80.94% Conservative: 51
Best Local Similarity: 70.49% Mismatches: 75
Query Match: 69.44% Indels: 18
DB: Gaps: 6
```

US-09-856-018b-16 (1-504) x AF251794 (1-1537)

```
Qy 21 ArgArgProSerAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThrAsp 40
Db 97 CGAAACGAGATCGGATCTTCCATGAGCTCTT-----CTGATTCGTAACTGTT 150
Qy 41 SerSerGlyAspLeuAlaLysAspSerGlySerAspAspSerIleAsnSerAspAsp 60
Db 151 TCCGATGCTGAC-----GTGAGGATCGG---GTTGATTCAGCTGTTGAGGATCTCAA 201
Qy 61 AlaAlaValAsnSerGlnGlnAsnGluLysGlnAspThrAspPheSerVal----- 78
Db 202 GGAACCAACCAATTTGGCCGAGAGAAACAAATTAGGAATCCGGTGGAGAGCGGGGA 261
Qy 79 -----LeuLysPheAlaTyrArgProSerValProAlaHisArgLysValLysGlu 95
Db 262 AACGTGGATGTAGGTACAGTATCGGCCGCTCGGTCGCTCAGCTCATCGAGGTCGGAG 321
Qy 96 SerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCys 115
Db 322 AGTCCACTCAGCTCAGCGCATCTTCAACAGAGCCATGCTGGACTATTCAACCTGTGT 381
Qy 116 IleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGly 135
Db 382 GTAGTAGTTCTTGTGCTGTAAACAGTAGACTCATCATCGAATAATCTCATCAAGTACGGT 441
Qy 136 TrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPhe 155
Db 442 TGGTTGATCAGACTGATTTCTGGTTAGTTCACGCTCTCGCAGGATTCGCCCTTTTC 501
Qy 156 MetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGluLysLeuAla 175
Db 502 ATGTGTTGCTCTCCCTTTCAATCTTTTCCCTTGGCTGCCCTTACCGTCGAGAAATAGTA 561
Qy 176 GlnArgLysCysIleProGluProValValValLeuHisIleIleIleThrSerThr 195
Db 562 CTTCAGAAATGCATATCTGAACCTGTTGTCTCATCTTTCATATATTATCACCATCACC 621
Qy 196 SerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheValSerGlyVal 215
Db 622 GAGGTCTGTATCCAGCTATGTGCTACTGACTCAAGGTGTGATTCGCCCTTCTTATCAGGTCTC 681
Qy 216 ThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsn 235
```

```
Db 682 ACCTTGATGCTCCTCCTACCTTCGATGTTGGTGGCGAAGTTGGTTCTTACCTCATATAAC 741
Qy 236 TyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeu 255
Db 742 TATGACATAAGAACCCTAGCTAATTCATCTGATAGGCC----- 780
Qy 256 AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaPro 275
Db 781 AATCTGAGTCTCCTCATATGTTAGCTGAAGAGCTTGGCTATTATGCTTGCTGCC 840
Qy 276 ThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPhe 295
Db 841 ACATTGTTTATCAGCGGAGCTATCCACGTTCTCCATGTATCCGGAAGGTTGGGTGGCT 900
Qy 296 ArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheIleIleAspGlnTyr 315
Db 901 CGTCAATTGCAAGCTGATCATATTCACCTGGATTCATGGGATTTAATAGAGCAATAT 960
Qy 316 IleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThr 335
Db 961 ATAAATCTTATTGTAGGAACCTCAAAACATCCTTTGAAAGGGATCTCTTATACGGTGT 1020
Qy 336 GluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCys 355
Db 1021 GAAACAGTGTGAAGCTTTCAGTTCCAAATTTATACGTGGCTCTGCATGTTCTACTGC 1080
Qy 356 PhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuLeuLeuLeuLeuLeuLeu 375
Db 1081 TTCTCCACCTTTGGTTAAACATATTGCGACAGCTCCTCTGCTCGGGGATCGTGAATTC 1140
Qy 376 TyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetPro 395
Db 1141 TACAAGATTGGTGAATGCAAAAAGCGTGGAGATATTATGGAGAAATGTGGAATATGCT 1200
Qy 396 ValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLys 415
Db 1201 GTTCATAAATGATGGTTCGACATGTATCTTCCGTGCTTCGCAGAAATATACCGAAA 1260
Qy 416 AlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAla 435
Db 1261 GPACCGCTATTATCTGCTTCTTAGTCTCTGCAGTCTTTCATGAGTTATGCATCGCA 1320
Qy 436 ValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeu 455
Db 1321 GTTCTTGTGCTGCTTCAAACTATGGGCTTCTTGGGATTTATGTTTCAGGTGCTTTG 1380
Qy 456 ValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIle 475
Db 1381 GTATTATCACAACTACCTACAGAAAGGTTT---GGCTCCATGGTGGGAAACATGATA 1437
Qy 476 PheTrpPheIlePheSerIleLeuGlnProMetCysValLeuLeuTyrTyrHisAsp 495
Db 1438 TTCTGGTTTACCTTCTGCAATTTTCGACAAACCGATGTGTGCTTCTTTATTATCAGCAC 1497
Qy 496 LeuMetAsnArgLysGlyLysLeu 503
Db 1498 TTGATGAACCGCAAGGAAGATG 1521

RESULT 4
LOCUS AY084052
DEFINITION Tropaeolum majus putative diacylglycerol acyltransferase mRNA,
complete cds.
ACCESSION AY084052
VERSION AY084052.1 GI:20086323
KEYWORDS
SOURCE Tropaeolum majus.
ORGANISM Tropaeolum majus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Tropaeolaceae; Tropaeolum.
REFERENCE
1 (bases 1 to 2090)
```

```
AUTHORS Mietkiewska,E., Pedersen,K., Katavic,V. and Taylor,D.C.
TITLE Characterization of a putative diacylglycerol acyltransferase mRNA
from Tropaeolum majus embryo
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2090)
AUTHORS Mietkiewska,E., Pedersen,K., Katavic,V. and Taylor,D.C.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2002) Seed Oil Biotechnology, National Research
Council of Canada, Plant Biotechnology Institute, 110 Gymnasium
Place, Saskatoon, SK S7N 0W9, Canada
FEATURES
source 1..2090
/organism="Tropaeolum majus"
/db_xref="taxon:4020"
171..1727
CDS
/codon_start=1
/product="putative diacylglycerol acyltransferase"
/protein_id="AA03340.1"
/db_xref="GI:20086324"
/translation="MAVAESQNTTMSGHGSDLNPNRRKRKSSSVIEPSSSGTST
NGVPATGHVAENRQDRVGAENATGSYNLIGNGGVVIGNEERQKVGEDIRFTYRPS
FPAHRRVRESPLSDAIFKQSHAGLNLIVLVANSRLIENLKMKGWLDYGFWF
NSRSLGWSIEMCCLTLPFFLAAFIVEKLVORNHISLVAVLLHVIYSTAALYPI
VILTCDSVYMSGVVLMFGCIMWLKLYSAHTSSDIPTLAKSGYKGDHPNSTIVSCS
YDLSKSLAYEWPVAPTLGQYSPSSRCIRKWVRQFVKLIVTIGLMGFIETQINP
IVRSKKHPLKGDLYAIERVLKLSPNDLYWLCMFYSFFHLMNLIALLELLRFGREFY
KDMNAKTVAEYWKWNNPVRHLYFPCLRNGIPKEGAIIFAFVSGFAHELCI
AVPCHVFKLWAFIGIMFQVPLVLTNTYLQERFSNWSVGNMIFWTFICILQPMCVLLY
YHDLINLKEK"
BASE COUNT 527 a 397 c 458 g 708 t
ORIGIN
Alignment Scores:
Pred. No.: 4.87e-153 Length: 2090
Score: 1836.00 Matches: 342
Percent Similarity: 78.05% Conservative: 67
Best Local Similarity: 65.27% Mismatches: 87
Query Match: 68.51% Indels: 28
DB: 8 Gaps: 5
US-09-856-018B-16 (1-504) x AY084052 (1-2090)
Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
Db 171 ATGGCGGTGGCAGATCGTGCAGAACACGACACACCATGAGTGGTCACGGCGACGCGAT 230
Qy 19 -----LeuArgArgProSerAla----- 25
Db 231 CTCACAATTTCCGTAGAGGAACCGAGTTCCTCCGTGATTGAACCTTCGTCGTCGGGT 290
Qy 26 ----ThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThrAspSerSerGlyAsp 44
Db 291 TTTACATCCACCAATGGGCTA-----CGGGCGACTGCCAGCTGAGAAATCGT 341
Qy 45 Asp-----LeuAlaLysAspSerGlySerAspAspSerIleAsnSer 58
Db 342 GACCAGGATCGGGTAGGGCTATGGAGAACGCAACAGGATCGGTCAACTTAATTGGAAT 401
Qy 59 AspAspAlaAlaValAsnSerGlnGlnAsnGluLysGlnAspThrAspPheSerVal 78
Db 402 GCTGGAGCGCTGTTATCGGGAATGAAGAAACAGGATAGGGAGACTGAT----- 452
Qy 79 LeuLysPheAlaTyrArgProSerValProAlaHisArgLysValLysGluSerProLeu 98
Db 453 ATACGATTCACTTACCGGCTTCGTTTCGGCTCATCGGAGGTGAGGAGAGTCTCTTT 512
Qy 99 SerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValVal 118
Db 513 ACCTCTGATGCAATCTTCAACAGAGCCCATCGGGTTTATTCAACTGTGTAGTAGTG 572
Qy 119 LeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTrpLeuIle 138
Db 573 CTCATTGCAGTAAACAGTAGGCTTATCATCGAAATCTTATCAAGTATGCTTGGTTGATC 632
```

```
Qy 139 LysSerGlyPheTrpPheSerLysSerLeuArgAspTrpProLeuPheMetCysCys 158
      :::::::::::::::::::::
Db 633 GATACTGTTCTGGTTAACTCAAGATCACTGGGTGATGGTCCATCTTATGTGCTGT 692

Qy 159 LeuSerLeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLys 178
      :::::::::::::::::::::
Db 693 CTTACACTCCCAATTTCCCACTTGCCTTTTATTTGAAAGAGCTGGTGACAGAAAT 752

Qy 179 CysIleProGluProValValValLeuHisIleIleIleTrpSerLeuPhe 198
      ||| ||| ||| :::::::::::::::::::::
Db 753 CATATATCTGAACCTGTTGCTTCTCCTTCATGTATCTTCTACCGCTGCAGTTTA 812

Qy 199 TyrProValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMet 218
      ||| ||| ||| :::::::::::::::::::::
Db 813 TATCCAGTATTATGTGATCTTAACGTGTGATTCGGTGATATGCTGCTGGTATTCATG 872

Qy 219 LeuPheSerCysValValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMet 238
      ||| ||| ||| :::::::::::::::::::::
Db 873 CTCCTTGGTTCGATATATGTTGAGCTGCTCATATGCACATAGTTCGTGATATT 932

Qy 239 ArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAsp 258
      ||| ||| ||| :::::::::::::::::::::
Db 933 AGAACACTGGCCCAATCTGGCTATAGGGGGATCGCCACCCCAATTCAACCATTTGAGT 992

Qy 259 TyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCys 278
      ||| ||| ||| :::::::::::::::::::::
Db 993 TGCTCATATGATGTAGTATGAAGAGTTGGCATCTTATGTTGGCGCCGACATTATGT 1052

Qy 279 TyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeu 298
      ||| ||| ||| :::::::::::::::::::::
Db 1053 TACCAGCTAGCTATCTGCTGCTGCTGATCCGCAAGGGTGGGTGTTGTCGCAATTT 1112

Qy 299 ValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnPro 318
      ||| ||| ||| :::::::::::::::::::::
Db 1113 GTCAACTATAGTATTTTCATAGGACTCATGGGTTTCATATAGAACCAATATATTAATCCT 1172

Qy 319 IleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgVal 338
      ||| ||| ||| :::::::::::::::::::::
Db 1173 ATCGTTGCAAAATTCCAACACCCCACTTCAAGAGAGATTTTATATGCAATAGAAAGAGTT 1232

Qy 339 LeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHis 358
      ||| ||| ||| :::::::::::::::::::::
Db 1233 TTGAAGCTTTTCAGTTCCAAATCTATATGTTGGCTTTGCGATTTCTACTCTTTTTCAC 1292

Qy 359 LeuTrpLeuAsnIleLeuAlaGluLeuArgPheGlyAspArgGluPheTyrLysAsp 378
      ||| ||| ||| :::::::::::::::::::::
Db 1293 CTCTGGTTGAACATACCTGGCTGAGCTTCTTCGCTTTGGTGATCGTGAATTTCTACAAAGAT 1352

Qy 379 TrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLys 398
      ||| ||| ||| :::::::::::::::::::::
Db 1353 TGSTGGAATGCAAAACTGTTGGGAGTATTTGGAATATGGAATATGCTGTTTCATAGA 1412

Qy 399 TrpMetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaAla 418
      ||| ||| ||| :::::::::::::::::::::
Db 1413 TGGATGTTGCTCATCTATATTTTCCCTGTTGAGGAATGGGATACCCCAAGGAGGTGCC 1472

Qy 419 LeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCys 438
      :::::::::::::::::::::
Db 1473 ATTATTTATCGGTTCTTAGTTTCTGGTGCTTTCATGAGCTCTGCATTCGATTCCTTGC 1532

Qy 439 HisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIle 458
      ||| ||| ||| :::::::::::::::::::::
Db 1533 CAGTATTCAAGTATTGGGCTTTATAGGCATTTATGTTTCAGGTTTCCTTGGTATTGATT 1592

Qy 459 ThrAsnTyrLeuGlnAsnLysPheArgAspSerMetValGlyAsnMetIlePheTrpPhe 478
      ||| ||| ||| :::::::::::::::::::::
Db 1593 ACCAATTTATCTACAAGAAAAGTTTCAGTAAATTTCTATGTTGGGCAATATGATCTTCTGCTTC 1652

Qy 479 IlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAspLeuMetAsn 498
      ||| ||| ||| :::::::::::::::::::::
Db 1653 ATCTTCTGCATACTTGGCCAACTATGTTGCTCTTCTATATACCATGACCTGATAAAT 1712,
```

```
Qy 499 ArgLysGlyLys 502
      ||| |||
Db 1713 CTRAAAGGAAAAG 1724

RESULT 5
AF051849 1845 bp mRNA linear PLN 24-JAN-2000
LOCUS Arabidopsis thaliana diacylglycerol acyltransferase (DAGAT) mRNA,
DEFINITION complete cds.
ACCESSION AF051849 AF061563
VERSION AF051849.1 GI:6625552
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Bouvier-Nave,P., Benveniste,P., Oelkers,P., Sturley,S.L. and
Schaller,H.
TITLE Expression in yeast and tobacco of plant cDNAs encoding acyl
CoA:diacylglycerol acyltransferase
JOURNAL Eur. J. Biochem. 267 (1), 85-96 (2000)
MEDLINE 20069349
PUBMED 10601854
REFERENCE 2 (bases 1 to 1845)
AUTHORS Benveniste,P.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-1998) Institut de Biologie Moleculaire des
Plantes, Centre National de la Recherche Scientifique, 28 rue
Goethe, Strasbourg 67083, France
COMMENT On Jan 24, 2000 this sequence version replaced gi:6650202.
FEATURES
Location/Qualifiers
source
1..1845
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/tissue_type="hypocotyl"
/clone_lib="ABRC library, catalog number CD4-15"
/db_stage="3 day-old seedling"
/notes="screened with EST, GenBank Accession Number
AA042298"
gene
1..1845
/genes="DAGAT"
88..1650
/genes="DAGAT"
/function="catalyzes the acylation of diacylglycerol by
oleoicoA to give triacylglycerol"
/notes="acyl-CoA cholesterol acyltransferase homolog"
/product="diacylglycerol acyltransferase"
/protein_id="AAF19362.1"
/db_xref="GI:6625553"
/translation="MAILDSAGVTTVTENGGEFVLDRLRLRRKSRSDSSNGLLLSGS
DNNPSDDVGAPADVDRIDSDVNDQGTANLAGDNGDNGGNGGGEGGRGNAD
ATFTPSVPAPHRARRSPSLDAIFKQSHAGLFLNCLVYVLIANSLRIENLMKYGW
LRTDFWSSRLSDWPLMCCISLIFPLAAFTVEKLVLQKYSVPVWIFLHIITM
TEVLYPVVYTLRCDSAFLSGVTLMLLTCIVLWKLVSVAHTSYDIRLSLANAADKANPEV
SYVLSKLAYFMVAPLTCYOPSPRSACIRKGWVARQFAKLVIFTGFPGFIEQIN
PIVRNSKHPKLGDLIIAERYLKLSPNLYVWLICMEYCFPHLMNLIAELLFCGDFE
YKDWNAKSPGDYRWMMNPVHKWVRHHYFPCLRSKIPKTLIIIIAFLVSVPFHELC
IIVPCLFLKFLWAFILGIMFQVPLVFTITNQLQERFGSTVGNMIFWIFCIFGQPMCVLLY
YHDLNMRKGSMS"
```

BASE COUNT 453 a 378 c 421 g 593 t

ORIGIN

Alignment Scores:

Pred. No.: 8,93e-152 Length: 1845

Score: 1821.00 Matches: 353

Percent Similarity: 75.33% Conservative: 47

Best Local Similarity: 66.48% Mismatches: 91

Query Match: 67.95% Indels: 40

DB: Gaps: 9

```
US-09-856-018B-16 (1-504) x AF051849 (1-1845)
QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
DB 88 ATGGCGATTGTTGATTCT--GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 144
QY 19 -----LeuArgArgArgProSer-----AlaThrSerThrAlaGlyLeu 31
DB 145 TTCTGCGATCTTGTATAGGCTTCGTGCGAGCAATCGAGATCGGATTTCTTCTAACGGACTT 204
QY 32 PheAsnSerProGluThrThrAspSerSerGlyAspLeuAlaLysAspSerGly 51
DB 205 CTTCTCTGTTGTCGATATAATCTCTCTCGGATGATGTTGGAGCTCCCGCGAGCTT 264
QY 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67
DB 265 AGGATCGGATTTGATTCCTGTTAAAGATGACGCTCAGGAAACAGCCCAATTTGGCCGGA 324
QY 68 GlnAsnGlu-----LysGln 72
DB 325 GATAATAACGGTGGTGGCGATATAACGGTGGTGGAGAGCGCGGAGAGGAGGA 384
QY 73 AspThrAspPheSerValLeuLysPheAlaTyArgProSerValProAlaHisArgLys 92
DB 385 ACGCGGATGCTAGC-----TTTACGTATCGACCGCTCGGTTCCAGCTCATCGGAG 435
QY 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112
DB 436 GCGAGAGAGAGTCCACTAGCTCCGAGCAATCTTCAACACAGAGCCATCCCGGATTATTC 495
QY 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleLeuAsnLeuMet 132
DB 496 AACCTCTGTGTAGTAGTCTTATGCTGTAAACAGTAGACTCATCATCGAAAATCTTAGT 555
QY 133 LysTyrglyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152
DB 556 AAGTATGGTGGTGTATCAGACAGGATTTCTGTTTGTAGTTCAAGATCGCTGGCAGATTGG 615
QY 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172
DB 616 CGCGTTTTCATGTGTATATCCCTTTCGATCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 675
QY 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValValLeuHisIleIle 192
DB 676 AAATGGTACTTTCAGAAATACATATCAGAACCTGTGTCTATCTTCTTCTATATTATTC 735
QY 193 ThrSerThrSerLeuPheTyProValLeuValIleLeuArgCysAspSerAlaPheVal 212
DB 736 ACCATCAGACAGGTTTGTATCCAGTTTACGTCACCTAAGGTGTGATTCGCTTTTATTA 795
QY 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyAla 232
DB 796 TCAGGTGTCACTTGTATGCTCTCACTTCATTTGTGGCTAAAGTTGGTTCTTATGCT 855
QY 233 HisThrAsnTyArgMetArgAlaLeuThrLysLeuValLcLysGlyGluAlaLeuLeu 252
DB 856 CATACTAGTATGACATAATGATCCCTAGCAATGACGAGTGAAGGCC----- 903
QY 253 AspThrLeuAsnMetAspTyProTyArgValSerPheLysSerLeuAlaTyPheLeu 272
DB 904 -----AATCCGGAAGTCTCCTACTACGTTAGCTTGAAGAGCTTGGCATATTTCATG 954
QY 273 ValAlaProThrLeuCysTyrglnProSerTyProArgThrProTyIleArgLysGly 292
DB 955 GTCGCTCCACATGTGTATCAGCAAGTTATCCAGCTTCTGCATGTATACGGAAGGCT 1014
QY 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312
DB 1015 TGGGTGGCTGCAATTTGCAAACTGGTCAATATTACCGGATTCATGGGATTTAATA 1074
QY 313 AspGlnTyIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
DB 1075 GAACAATATATAAATCCTATTGTCAGGAACCTCAAGCATCTTCTTGAAGGCCATCTTCTA 1134
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyValTrpLeuCysMet 352
DB 1135 TATGCTATTGAAAGAGTGTGAAGCTTTTCAAGTTTCAAAATTTATATGTGTGGCTCTGCATG 1194
QY 353 PheTyrcysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
DB 1195 TTCTACTGCTTCTTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTCTGCTTCGGGAT 1254
QY 373 ArgGluPheTyrcysAspTrpAsnAlaLysThrValGluAspTyTrpArgMetTrp 392
DB 1255 CQTGATTTCTACAAAGATTTGGTGAATGCAAAAGTGTGGGAGATTACTGGGAATGTGG 1314
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrcysLeuArgHisGly 412
DB 1315 AATATGCTGTTTCATAAATGGATGTTTCGACATATATATACTTCCCGTGTCTGCGCAAG 1374
QY 413 LeuProLysAlaAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
DB 1375 ATACCAAAACACTCGCCATTATCATTTGCTTCTAGTCTCGAGTCTTTCATGAACTA 1434
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyIleMetPheGln 452
DB 1435 TGCATCGCATGTTCTTGTGCTCTCTTCAACCTATGGGCTTTTCTTGGGATTTATGTTTCAG 1494
QY 453 ValProLeuValLeuIleThrAsnTyrcysLeuGlnAsnLysPheArgAsnSerMetValGly 472
DB 1495 GTGCTTGTGCTTTCATCAACAACTATCTACAGGAAGGTTT---GGCTCAACGGTGGG 1551
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeuTyrc 492
DB 1552 AACATGATCTTCTGTTTCATCTTCTGCAATTTTCGGCAACCGATGTGTGCTTCTTTAT 1611
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
DB 1612 TACCAGCACTGATGACCCGAAAGGATCGATG 1644
RESULT 6
ATH238008 1904 bp mRNA linear PLN 18-JUN-1999
LOCUS Arabidopsis thaliana mRNA for diacylglycerol acyltransferase.
DEFINITION Arabidopsis thaliana mRNA for diacylglycerol acyltransferase.
ACCESSION AJ238008
VERSION AJ238008.1 GI:5123717
KEYWORDS Dgat gene; diacylglycerol acyltransferase.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1904)
AUTHORS Zou,J., Wei,Y., Jiao,C., Selvaraj,G. and Taylor,D.C.
TITLE The Arabidopsis thaliana TAG1 gene encodes for a diacylglycerol
JOURNAL acyltransferase
REFERENCE 2 (bases 1 to 1904)
AUTHORS Zou,J.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1999) Zou J., Plant Biotechnology Institute,
National Research Council of Canada, S7H 0W9, Saskatchewan, CANADA
FEATURES
source
1. .1904
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
139. .1701
/gene="Dgat"
139. .1701
/gene="Dgat"
/function="lipid synthesis"
/codon_start=1
/product="diacylglycerol acyltransferase"
/protein_id="CAB45373.1"
/db_xref="GI:5123718"
```


[illegible]

Oilseeds Research, John Innes Centre, Colney Lane, Norwich, NR4 7UH, U.K

FEATURES
source Location/Qualifiers
1..1988 /organism="Arabidopsis thaliana"
/sub_species="Columbia"
/db_xref="taxon:3702"
CDS 231..1793
/EC_number="2.3.1.20"
/codon_start=1
/product="diacylglycerol O-acyltransferase"
/protein_id="CAB44774.1"
/db_xref="GI:5050913"
/translation="MALIDSAGVTVTENGGEFVDLRLRRKRSRSDSSNGLLLSGS
DNPSDDVGAPADVRIDSVVNDQAGTANLAGDNNNGGNGGGGGGGEGGEGGNAD
ATFTRFSPAHRRARSPPLSSDAIFKQSHAGLFNLCVVVLIAYNSRIIENLMKYGW
LIRDFRFSRSLRDWPLMCCISLSPFLAAFTVEKVLQKYSIEPVVIFLHIITM
TEVLYPVVTVLRCDASFLSGVTMLLTCTIVMLKLVSYAHTSVDRSLANAADKANPEV
SYVYSKSLAYFMVAPTLCYOPSPRSACIRKGMVAROFKLVIFGTGPMGFTLIBQYIN
PIVNSKPLAGDLLIALERVLKLSVPLNYVLMCMFYCFPHLNLIAELLCLFGDREF
YKDMNNAKSGDIWRMMNMPVHKWVRHIYFPCLSKSLPKTLALIIIAFLVSAVFELC
IATVPCRLFKLWAFIGMFQVPLFITNYLQERFSGTVGNMIFWIFCFIFGQPMCVLLY
YHDLNRRKGSN"

BASE COUNT 478 a 423 c 436 g 651 t
ORIGIN

Alignment Scores:
Pred. No.: 9.77e-152 Length: 1988
Score: 1821.00 Matches: 353
Percent Similarity: 75.33% Mismatches: 47
Best Local Similarity: 66.48% Indels: 40
Query Match: 67.95% Gaps: 9
DB: 8

US-09-856-018B-16 (1-504) x ATH131831 (1-1988)

Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
Db 231 ATGGCGATTGTTGGATTCT---GCTGGCGTTACTAGCGTGACGGAACGGTGGCGGAGAG 287
Qy 19 -----LeuArgArgProSer---AlaThrSerThrAlaGlyLeu 31
Db 288 TTCGTCGATCTTGATAGGCTTCGTCGACGGAATCGAGATCGGATCTCTTAACGGACTT 347
Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspPheAlaLysAspSerGly 51
Db 348 CTTCTCTCTGTCGATAATAATCTCTCTCGGATGATGTTGGAGCTCCCGCGACGTT 407
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67
Db 408 AGGGATCGGATTGATTCGCTTTAACGATGACGCTCAGGGAACAGCCCAATTTGGCCGGA 467
Qy 68 GlnAsnGlu-----LysGln 72
Db 468 GATAATAACGGTGGTGGCCATAATAACGGTGGTGGAGAGGCGCGGAGAGGAGA 527
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyArgProSerValProAlaHisArgLys 92
Db 528 AACGCCGATGCTACG-----TTTACGATACGACGCTCGGTTCCAGCTCATCGGAGG 578
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112
Db 579 GCGAGAGAGATTCCTAGCTCCGACGCAATCTTCAACAGAGCGCATCGCGGATTATTC 638
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132
Db 639 AACCTCTGTAGTAGTCTTATTCTGTAACACAGTAGACTCATCATCGAAATCTTATG 698
Qy 133 LysTyGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152
Db 699 AAGTATGTTGTTGATCAGAACGGATTCTCGTTTAGTTCAAGATCGTGGCAGATTGG 758
Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172

Db 759 CCGCTTTTCATGTTGTATATATCCCTTTCCATGCTTTCCCTTTGGCTGCCCTTTACGGTTGAG 818
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValValLeuIleIleIle 192
Db 819 AAATGGTACTTCAGAAAATACATATACAGAACCTGTTGTCATCTTCTTCATATATTATTC 878
Qy 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
Db 879 ACCATGACAGAGGTTTGTATCCAGTTTACCTACCCCTAACGTTGATCTCTCTTTT 938
Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232
Db 939 TCAGGTGTCACCTTGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 998
Qy 233 HisThrAsnTyAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
Db 999 CATACTAGTATGACATAAGATCCCTAGCCATGACGAGCTGATAAGGCC----- 1046
Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyPheLeu 272
Db 1047 -----AATCCTGAAGTCTCTCTACTACTAGCTTAGCTTGAAGAGCTTGGCATATTTTCATG 1097
Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProAspThrProTyrIleArgLysGly 292
Db 1098 GTCCGCTCCACATTTGTTTATCAGCCAGTTATCCACGTTCTGCGATGTATACGGAAGGT 1157
Qy 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312
Db 1158 TGGGTGGCTCGTCAATTTGCCAAACTGGTCATATTCACCGGATTCATGGGATTTATAATA 1217
Qy 313 AspGlnTyIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
Db 1218 GAACAATATATAATCTCTATTCAGGAACCTCAAGAGCTCTTTGAAGCGCATCTCTCA 1277
Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
Db 1278 TATGCTATTGAAAGAGTGTGAAGCTTTTCAGTTCCAAATTTATATGTGTGTGCTGTCATG 1337
Qy 353 PheTyCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
Db 1338 TTCTACTGCTTCTCCACCTTGGTTAAACATATTTGGCAGAGCTTCTCTGCTTCGGGAT 1397
Qy 373 ArgGluPheTyLysAspTrpTrpAsnAlaLysThrValGluAspTrpTrpArgMetTrp 392
Db 1398 CGTGAATTCACAAAGATTGGTGAATGCAAAAGTGGGAGATTACTGGAGATGTGG 1457
Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyPheProCysLeuArgHisGly 412
Db 1458 AATATGCCCTGTTCAAAATGGATGGTTCGACATATATATCTCCCGTGTTCGCGACAG 1517
Qy 413 LeuProLysAlaAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1518 ATACAAAGACACTGCCCATTTATCATGTTTCCTAGTCTCTGACGCTTTCATGAGCTA 1577
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyIleMetPheGln 452
Db 1578 TGCATCGCAGTTCCTTGTGCTCTCTCAAGCTATGGCTTTCTTGGGATTTATGTTTCAG 1637
Qy 453 ValProLeuValIleThrAsnTyLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1638 GTGCCCTTGGTCTTCATCAACAATCTACAGGAAGAGTTT---GGCTCAACGGTGGG 1694
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
Db 1695 AACATGATCTCTGGTTCATCTTCTGCAATTTTCGACAAACCGATGTGTGTCTCTTAT 1754
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1755 TACCACGACCTGATGAACCGAAAGGATCGATG 1787
RESULT 10
AT05480


```
QY 313 AspGlnTyrIleAsnProfilValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
Db 1149 GAACAATATATAAATCCTATTCTCAGGAAGTCAAGAGCCTCTTTGAAAGCGCATCTCTTA 1208
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
Db 1209 TATGCTATTGAAGAGTGTGAAGCTTTCAGTTCCTCCAAATTTATATGTGTGGCTCGCATG 1268
QY 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
Db 1269 TTTCTACTGCTTCTCCACCTTTGGTAAACATATTGGCAGAGCTTCTCTGCTTCGGGAT 1328
QY 373 ArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
Db 1329 CQTGAATCTCACAAGATTGGTGAATGCAAAAGTGTGGGAGATTACGTGGAGATGTGG 1388
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
Db 1389 NATATGCTGTTTCATAAATGGATGCTTCGACATATATATCTTCCGCTGCTGGCAGCAG 1448
QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1449 ATPACAAAGACACTCGCCATTATCATGCTTCTTCTAGTCTCTGCACTCTTTCATGAGT 1508
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyClyIleMetPheGln 452
Db 1509 TGCATCGCAGTTCCTGTGCTCTCTCAAGCTATGGCTTTCTTGGGATTATGTTTCAG 1568
QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1569 GTGCTTTGGTCTTCATCACAACATATCTACAGGAAGGTTT--GGCTCAACGGTGGG 1625
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
Db 1626 AACATGATCTTCTGCTGTCATCTCTGCAATTTTCGACACACCATGCTGTGCTTCTTAT 1685
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1686 TACCAGCCTGATGACCGAAAGATGCGATG 1718

RESULT 11
AF164434 1512 bp mRNA linear PLN 30-NOV-1999
LOCUS Brassica napus putative diacylglycerol acyltransferase (DGAT1)
DEFINITION mRNA, complete cds.
ACCESSION AF164434
VERSION AF164434.1 GI:5579407
KEYWORDS
SOURCE Brassica napus.
ORGANISM Brassica napus.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
AUTHORS Nykiforuk,C.L., Laroche,A. and Weselake,R.J.
TITLE Isolation and Characterization of a cDNA Encoding a Second Putative
Diacylglycerol Acyltransferase from a Microspore-derived Cell
Suspension Culture of Brassica napus L. cv Jet Neuf (Accession No.
AF164434). (PCR99-158)
JOURNAL Plant Physiol. 121 (3), 1053 (1999)
REFERENCE
AUTHORS Nykiforuk,C.L., Laroche,A. and Weselake,R.J.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Chemistry and Biochemistry, University of
Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4,
Canada
FEATURES
Source Location/Qualifiers
1..1512
/organism="Brassica napus"
/strain="cultivar Jet Neuf"
/db_xref="taxon:3708"
/note="microspore-derived cell suspension culture"
1..1512
gene
```

```

/ gene="DGAT1"
1..1512
/ gene="DGAT1"
/ EC_number="2.3.1.20"
/ codon_start=1
/ product="putative diacylglycerol acyltransferase"
/ protein_id="AAD45536.1"
/ db_xref="GI:5579408"
/ translation="MAILDSGGVAVPPTENGVAIDLRLHRRKSSSDSSNGLLSDTSPS
DDVGAARERDRVDSAAEEAEAGTANLAGDAETRESAGGVRFRTYRPSVPAHRRTRE
SPLSDDAIFKQSHAGLFNLVAVASRLIENLMKYGLWIRTFDFWFSSTSLRDWP
LFMCCLSLSPFLAAFTVEKVLQKFISEPVAILHLVITTEVLVYPVTVLRCDSEF
LSGYTMLLTCTVWLKLVSYAHTSDIRTLANSADKVDPEISYVLSKSLAYFWVAPT
LCYQPSYPRSPCLRWARGWAKLAVIFTEGLMGFTIEQIIMPVNSRHPKLGDDLLYA
IERVLKSVPLNYLWLCMFYCFHLLNLAEELCFGDREYFKDMNASKVDGYWRMW
NPVHKMWRHVYFCLRIKIPKVAIIIAFLVSAVFEHLCTIAPVCRFLNFWAPMGIM
FOVPLVFTTNFLQERFSGVMGNIFGSASCIFQPMCGLLLYVHDLNRKGSMS"
BASE COUNT 344 a 343 c 363 g 462 t
ORIGIN
Alignment Scores:
Pred. No.: 1,56e-149 Length: 1512
Score: 1794.50 Matches: 335
Percent Similarity: 79.11% Conservative: 55
Best Local Similarity: 67.95% Mismatches: 78
Query Match: 66.96% Indels: 25
DB: 8 Gaps: 5
US-09-856-018B-16 (1-504) x AF164434 (1-1512)
QY 20 ArgArgArgProSerAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThr 39
Db 76 CQCTGTAATCAGTTCGGATTCTTCCAAACGGACTCTCTCCGAT----- 120
QY 40 AspSerSerGlyAspLeuAlaLysAspSerGlySerAspAspSerIleAsnSerAsp 59
Db 121 ACTTCCCGTCGACGATGTTGGAGTCGCGCGCGCGAAGGATCGGGT----- 171
QY 60 AspAlaAlaValAsnSerGlnGlnAsn----- 70
Db 172 GATTCGCTGCGAGGAGGAGGCTCAGGACAGCGAAATTTAGTCGCGGAGATGCCAA 231
QY 71 LysGlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHis 90
Db 232 ACTAGGAATCCCGCGAGGCGGATGTAAGTTTACGTATCGACCGCGGTTCAGGTCAT 291
QY 91 ArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGly 110
Db 292 CGGAGGACGAGGAGAGTCTCTCAGCTCCGACGCTATCTTCAACAAGCCATGCAGGA 351
QY 111 LeuPheAsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleLeuGluAsn 130
Db 352 TTGTTCAACCTCTGTGTAGTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
QY 131 LeuMetLysTyrClyTrpLeuLysSerGlyPheTrpPheSerSerLysSerLeuArg 150
Db 412 CTCATGAAGATGTTGGTGTGTATCAGAACTGATTTTGTGTGTGTGTGTGTGTGTGTGT 471
QY 151 AspTrpProLeuPheMetCysLysLeuSerLeuValValPheProPheAlaAlaPheIle 170
Db 472 GACTGGCGCTTTTCATGTTGTTCTTTCACCTTCGGTCTTTCCTTGGCTGCCTTCAGC 531
QY 171 ValGluLysLeuAlaGlnArgLysCysIleProGluProValValValLeuHisIle 190
Db 532 GTGAGAAAATGCTACTTCTAGAAATTCATATCTGAGCCCTGTGGCCATCATCTTCTATG 591
QY 191 IleIleThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAla 210
Db 592 ATTATAACCATGACAGAGGCTTGTGTATCCAGTCTACGTACAGTGTGATTCTGTC 651
QY 211 PheValSerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSer 230
Db 652 TTCTGTGTCAGGTGTACAGTGTGATGCTGCTCATTGCTGCTGCTGCTGCTGCTGCTTCT 711
```


Qy	295	heArgGlnLeuValLysLeuIlellePheThrGlyValMetGlyPheIleleasp----	313
Db	22722	CTCGTCAATTTGCAAAACTGTCATATTCACCGGATTCATGGGATTTATATAGAACAAG	22781
Qy	313	-----	313
Db	22782	TACGTTTTTCACATCTTGCTTTTATTAGTTTTCTTGGTGGAAATCATCATCCCTCGGTGT	22841
Qy	314	-----GlnTyrlleAsnProilleValGln	321
Db	22842	CACCACCTTGACTTCATGTTCTTTTGTATACATTTTGGCAGTATATAAATCCATTGTGCAGG	22901
Qy	322	AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrlalArhGluArGValLeuLysLeu	341
Db	22902	AACTCAAAAGCATCCCTTTGAAAGGGCATCTTCATATGCTATTGAAGAGTGTGAAGCTT	22961
Qy	342	SerValProAsnLeuTyrlValTrpLeuCySmetPheTyCySPhPheHisLeuTrp----	360
Db	22962	TCAGTTCCAAATTTATATGTGTGGCTCTGCATGTTCTACTGCTTCTCCACCTTTGGTAT	23021
Qy	360	-----	360
Db	23022	GCTGTGATCCATCTCTTTTCAAAATAATTGCAAAATTCGAAAAACCGAAAAAGCGCTAAAT	23081
Qy	360	-----	360
Db	23082	CTCATACGAATTTGTATATTTTAGTCTTCTAGAGTCGGTGATGTAATTTCTAGTTACTGAA	23141
Qy	361	-----LeuAsnIleLeuAlaGluLeuLeuArGPhGlyAspArG	373
Db	23142	CGCAATCTCTGTGCCAAAGGTTAAACATATGGCAGAGCTCTCTGCTTCGGGGATCGT	23201
Qy	374	GluPheTyrlLysAspTrpTrpAsnAlalysThrValGlu-----	386
Db	23202	GAATCTTCAAAAGATTGTGTGAATGCAAAAAGTGTGGG-AGATGTGAGCTATTTTACTCA	23260
Qy	386	-----	386
Db	23261	AAAGAAAACTTATGATTTTTTAATGTGTGCTGTTGTTTTGGTGCATCTAACTAACCAAAAT	23320
Qy	387	-----AspTyrlrPargMetTrpAsnMet----	394
Db	23321	CATGTATTCAGTGTCTTCCCTTTATCAGTACTGGGAATGTGGAATATAGGTATGTTCTCT	23380
Qy	394	-----	394
Db	23381	TCCTAAACATCACCTTCTTTGTACACAAATAGNAGAAGAGAGCTAATTAAGATCTTGT	23440
Qy	395	-----ProValHisLysTrpMetIleArgHisLeuTyrlPheProCySLeuArg	410
Db	23441	TTTCCTTGACAGCCTGTTTCATAAAATGGATGGTTGACATATATACTTCCCGTGTGCGC	23500
Qy	411	HisGlyLeuProLysAlaAla-----	417
Db	23501	AGCAAGATACCAAAAGGTGAG-TGAGATATATACCGATATACCGATATGCAATGTGAGATTGTTTC	23559
Qy	418	-----AlaLeuLeuIleAla	422
Db	23560	TGTGATATAAATTTAAACCTCCACACACTTGTTTTTCAGACACTCGCCATTCATTTGCT	23619
Qy	423	PheLeuValSerAlaLeuPheHisGlu-----	431
Db	23620	TTCTAGTCTCTGCAGTCTTTTCATGAGGTATACATACTTTCTACATTTGCCCTGTCTCTAG	23679
Qy	431	-----	431
Db	23680	ACGCATGAACACACGCTAGTGAAAGAAATGCTATATTTCAAAGCATGTTTTTACTTAAC	23739
Qy	432	-----LeuCysIleAlaValProCySHisIlePhe	441
Db	23740	GATCTGTGTTACAAATTTCCCTTTTGACAGCTATGCATCGCAGTTCCTTGTGCTCTCTTC	23799
Qy	442	LysLeuTrpAlaPheGlyGlyIleMetPheGln-----	452

```
|||||
Db 23800 AAGCTATGGCTTTTCTTGGGATTATGTTTCAGGTTAAAAAATTACTAAACTGCTGCAGT 23859
Qy 453 -----ValPro 454
|||||
Db 23860 CGATTTTAAACTCTAATCTCATATTCGACCAACCAATTTGTTGAGTAGGTGCCT 23919
Qy 455 LeuValLeuIleThrAsnTyrLeuGlnAsnLysPhe--ArgAsnSerMet----- 470
|||||
Db 23920 TTGGTCTTCATCACAAACTATCTACAGAAAGTTTGGCTCAACGGTATGCTCTCAAAAC 23979
Qy 470 ----- 470
|||||
Db 23980 CCGAGAAAATAGAACGAATAACTCTTCTTCATAGCCTAGCCATTTAAATCGCAATGCT 24039
Qy 471 -----ValGlyAsnM 474
|||||
Db 24040 GAAACTTAATAAAGGTGATCTGTTTGAATGGGATCATATATTAGTGGGGAACA 24099
Qy 474 ettlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeuTyrTyrH 494
|||||
Db 24100 TGATCTTCGTCTTCATCTTCGCAATTTTCGGACAAACCGATGTGTGCTTCTTATTACC 24159
Qy 494 IsAspLeuMetAsnArgLysGlyLysLeu 503
|||||
Db 24160 ACGACTGATGAACGAAAGATCGATG 24188
|||||
RESULT 14
AC003058 97495 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F27F23 map C1C06E08,
DEFINITION complete sequence.
ACCESSION AC003058
VERSION AC003058.3 GI:20197007
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 97495)
Rounsley,S.D., Lin,X., Ketchum,K.A., Crosby,M.L., Brandon,R.C.,
Sykes,S.M., Kaul,S., Mason,T.M., Kerlavage,A.R., Adams,M.D.,
Somerville,C.R. and Venter,J.C.
Unpublished
2 (bases 1 to 97495)
Lin,X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 97495)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@igir.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598388.
FEATURES
source
1. 97495
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C06E08"
/clone="F27F23"
1. 9642
/notes="overlap with BAC clone T20K24
(AC002392:95887..105528).
complement(21..3689)
/gene="At2g19210"
/note="F27F23.1"
complement(join(<21..429,506..695,770..832,927..1053,
1129..1340,1428..1594,1710..1790,1882..1953,2038..2109,
2186..2324,2406..2884,2963..3509,3602..3689))
misc_feature
1. 9642
/notes="overlap with BAC clone T20K24
(AC002392:95887..105528).
complement(21..3689)
/gene="At2g19210"
/note="F27F23.1"
complement(join(<21..429,506..695,770..832,927..1053,
1129..1340,1428..1594,1710..1790,1882..1953,2038..2109,
2186..2324,2406..2884,2963..3509,3602..3689))
CDS
complement(join(21..429,506..695,770..832,927..1053,
1129..1340,1428..1594,1710..1790,1882..1953,2038..2109,
2186..2324,2406..2884,2963..3509,3602..3689))
/gene="At2g19210"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC16451.1"
/db_xref="GI:3135251"
/translation="MVHYNFLSLIIFACFFAVFVLVLAQDSGFVSIDCGIPEDSSY
NDETDTIKYSDAAAFVESGIIHSIDPEFQSLEKQFQNVRSFPEGNNKCYDVKPPQG
KGKYLIRTRFMYGNLDNKGAPDFDLYLGFNIWDSVTIDNATITVTKRISDH
HVCLVDKNGRTPELSALEIRLLKSNITETPYDSLILFKRMDLGGALGALVYRKKDYF
DRWIPLRFPKTYIFNASLIDSNNGNEQFARFVMMNTASPEDLSODIIFSNPKDP
TWKYFYVMEFAVEVPELSPNETREFKVLNKEIKNMSFSRPLYVTDLTFLVQNPVSGPK
LEFRLOQTPRSTLPIIINATETRYVNEFLOSPTDODDVAIMRIKSKYGVKSKLGD
CAPVKPKDINGSYVDNESPRIISVSSGSLTGEIDAFAFSLTLHLIIDLNSST
GKIPDFGNHNLTELNEGNKUSGAIPKVLERSNKKLILRLIDGPNVAGPLDVCASOI
SDEKTKNVIYIPLVASVGVGLVLAIALFLYKKRRGGSGGVRAGPLDVTTRY
KYSEVVKVTNNFVRLVGGQGVYHGVLDNDQVAVKILSESAQGYKEFRAEVELLL
RVHKNLTALIGYCHEGKMAIYEFMANGTLGDYLSGKSYVLSWEERLQISLDAAG
GLEVLHNGCKPPIVORDVKPANTILINEKLOAKIADFGLSRVALDGNNOPTTAVAGTI
GYLDPVHLTKLSEKSDIYSFGVVLLEVYSGQPVIAISRTTAENIHTIDRVDLMLST
GDINGVDPKIGERFDSAWKITEVAMACASSSSKNRPTMSHVVAELKELSVSRARAG
GGSGASSVDPAMTNFDSGMFPQAR"
complement(5018..5047)
/rpt_family="At_rich"
complement(5060..6543)
/note="F27F23.2; predicted by genscan"
complement(join(<5060..5590,5665..6171,6262..6543))
/gene="At2g19220"
complement(join(5060..5590,5665..6171,6262..6543))
/gene="At2g19220"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC16452.1"
/db_xref="GI:3135252"
/translation="MTSKAAWEPEHDEVFVLDLQVEQKMLGNQEMOHILEAFOEMGYR
FTDQLINHWDTMIKQKWCVRQCKDKIWDLSLNTFGATDOEWANLYVPEAGQY
RNPPLLEKLEIIFAGMNDLGGTSGSKMKOICEBDEENVTVYVPRISADIAIR
RHYKWPSSHAIVVDCFOBSLAGIRPIKRNHLFTKESWKKILEKINRTGLGTYTHQ
LEHNFTRTISWAKWCETIASPIKMDANTARKFGATEEDWDKLYLMINKRNVPRKRRI
SDDRAEDGNVETSLRISDEDAETVPEMEKIPVNASVKKKEKTFEECEVCLDA
IEEVEKGDLYMEALDFTKDYRYFLMLQKSLRMAWLLKLSNP"
complement(5603..5623)
/rpt_family="At_rich"
complement(9141..12813)
/gene="At2g19230"
/note="F27F23.3"
complement(join(<9141..9558,9652..9841,9925..9987,
10073..10199,10282..10478,10565..10722,10838..10918,
11004..11075,11166..11237,11314..11452,11535..12019,
12091..12637,12729..12813))
/gene="At2g19230"
complement(join(9141..9558,9652..9841,9925..9987,
10073..10199,10282..10478,10565..10722,10838..10918,
11004..11075,11166..11237,11314..11452,11535..12019,
12091..12637,12729..12813))
/gene="At2g19230"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC16453.1"
/db_xref="GI:3135253"
/translation="MGNFNPPLVPSFASVFWVLVLAQDSGFVSIDCGIPEDSSYY
DEKTDIKYISDAAFVESGIIHSIDSKFQKNLEKQFQNVRSFPEGNNKCYDVKPPQK
GKYLIRTRFMYGNLDNKGAPDFDLYLGVNLWDSVTLENTITVTKRISDH
HVCLVDKNGRTPELSALEIRLLKSNITETPYDSLILFKRMDLGGALGALVYRKKDYF
RWMLPMFNFNLITSLMIDPTSSNGFTPPSVVMSTAVRAMNSISLTPARVYKEDIDF
NWKFYIIFHAVEKLPSPNETREFSVFLNKEQIDTTSVSPRSYLYTDLTVQNPVSGP
LEFLVLRQGVKSTPPIIINATETRYVNEFLOSPTDODDVAIMRIKSKYGVKSKLGD
PCAPGYPMQVINCSTIANNPPIIISVNLISGLTGIDQIDVPFITLPLQLKLDLSNNHL
```


Qy 157 -----CysCysLeuSe 160
Db 91586 ATAATGATGAGTCTAGTGAATCAAGTTCCTCTTTGATTCTCTCTTCTGTCAGTATATC 91645
Qy 160 rLeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180
Db 91646 CTTTCGATCTTCTCTTGGCTGGCTTTACGGTTGAGAAATTTGGTACTTCAAGAAATACAT 91705
Qy 180 eProGluProVal-ValValVal----- 187
Db 91706 ATCAGAACCTGTGAGTAATTAATCTCCAGCCATTACTGTAATTTTATTGAAGACAA 91765
Qy 188 -----L 188
Db 91766 GTTTGTATCATGAAGAACTTACAAGTTCTGTGTTTGAAGAAATGCTCAAGGTTGCTCATCTTTC 91825
Qy 188 euHisIleIleIleThrSerLeuPheTyrProValLeuValIleLeu----- 205
Db 91826 TTCAATATTTATCACCATGACAGAGGTTTGTATCCAGTTTACGTCACCCCTAAGGTGAT 91885
Qy 205 ----- 205
Db 91886 ACTGTTTCTGCTCAGTTTGTGATACCTGTTTTTAAGTTTACTGTCTGACCCGGTGA 91945
Qy 206 -----ArgCysaspSerAlaPheValSerGlyValThrLeuMetLeuPheS 221
Db 91946 TCTTCAAAATGGACAGGTGATGCTGTTTTTATCAGGTGTCACTTTGATGCTCTCTCA 92005
Qy 221 erCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaL 241
Db 92006 CTTGCATTTGTGGCTAAAGTTGGTTCTTATGCTCATCTACTAGCTATGACATAAGATCCC 92065
Qy 241 euThrLysLeuValGluLys-----GlyGluAlaLeuLeuAspThr 254
Db 92066 TAGCAATGTCAGCTGATAAGGTAAATACGAAAGAGCGTATGTTAT-AGTCATTCG 92124
Qy 255 -----LeuAsnMetAspTyrProT 261
Db 92125 ACTGTGTACTGTTTTTAACCAACACACTGTTATGAACTTTATAGGCCAATCCTGAAGTCTCT 92184
Qy 261 yrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyr----- 279
Db 92185 ACTAGTTAGTTGAAGAGCTTGCAATTTTCATGCTGCTGCCACATTTGTTATCAGG 92244
Qy 279 ----- 279
Db 92245 TAACTGCAAGTGCATCAACCATTTCTTATACTTGCAGAGTTTCTGTCTAAACCTCGGA 92304
Qy 280 -----GlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuP 295
Db 92305 TCTTTTGTCTTCCCGCCCAAGTTTATCCAGCTTCTGCATGTATACGGAAGGTTGGGTGG 92364
Qy 295 heArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAsp----- 313
Db 92365 CTCTGCAATTTGCAAACTGTCATATTCACCGATTCATGGGATTTATATAAGAACAG 92424
Qy 313 ----- 313
Db 92425 TACGTTTTACATCTTGTGTTTATTAGTTTCTTGTGTTGAAATCATCATCCCTCGGTTGT 92484
Qy 314 -----GlnTyrIleAsnProIleValGln 321
Db 92485 CACCACCTTGACTTCATGTTCTTTTGTACATTTTGGCAGTATATAAAATCTCTATTGTGAG 92544
Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeu 341
Db 92545 AACTCAAGCATCTCTTGAAGGCGATCTCTATATGCTATTGAAAGAGTTTGAAGCTT 92604
Qy 342 SerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrp----- 360
Db 92605 TCAGTTCCAAATTTATATGCTGGCTCTGCATGTTCTACTGCTCTCTCCACCTTTGGTAT 92664
Qy 360 ----- 360

Db 92665 GCTGTGATCCCATCTCTTTCAAAATAATTTGCAAAATTCGAAAAACCGAAAAAGGCTAAAT 92724
Qy 360 ----- 360
Db 92725 CTCATACGAATTTGATATTTTGTAGTCTCGGTGATGTAATTTTCTAGTTACTGAA 92784
Qy 361 -----LeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArg 373
Db 92785 CGCAAACTCTCTTGTCCAAAGGTTAAACATATTGCGACAGCTTCTCTGCTCGGGGATCGT 92844
Qy 374 GluPheTyrLysAspTrpTrpAsnAlaLysThrValGlu----- 386
Db 92845 GAATTCACAAAGATTTGGTGAATTCGAAATTCGAAAGTGTGGG-AGATGTGAGCTATTTTACTCA 92903
Qy 386 ----- 386
Db 92904 AAAGAAAACATTATGATTTTAAATGTTGCTGTTGTTTTGGTGCATCTTAACTAACCAAAAT 92963
Qy 387 -----AspTyrTrpArgMetTrpAsnMet----- 394
Db 92964 CATGTATTCACTGCTCTCTTTTATCAGTACTGAGAAATGTGGAATATGTTATGTTCTCT 93023
Qy 394 ----- 394
Db 93024 TCCTAAACATCACCTTCTTTGTACAAAAATAGAAGAGAGACTAATTAAAGATCTTGT 93083
Qy 395 -----ProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArg 410
Db 93084 TTTCTCTGACAGCTGTTCATAAATGGATGTTTCGACATATATATCTTCCCGTGTTCGCG 93143
Qy 411 HisGlyLeuProLysAlaAla----- 417
Db 93144 AGCAAGATACCAAGGTGAG-TGAGATATATACCGATATGCAATTTGTCGAGATTTGTTTC 93202
Qy 418 -----AlaLeuLeuIleAla 422
Db 93203 TGTGATATAATTTAACCTTCCACACACTTGTGTTTTTTCAGACACTCGCCATTTATCATGCT 93262
Qy 423 PheLeuValSerAlaLeuPheHisGlu----- 431
Db 93263 TTCTCTAGCTCTGAGCTTTTCATGAGGTATACATACTTTCTACATTTGCCCTGTCTAG 93322
Qy 431 ----- 431
Db 93323 ACGCATGAACACACACGCTAGTGAAGAAATGCTAATATTCAAAGCATTTGTTTTACTTAAC 93382
Qy 432 -----LeuCysIleAlaValProCysHisIlePhe 441
Db 93383 GATCTGTGTACAAATTTCTTTTGACAGCTATGTCATCGCAGTTCTCTTGTCTCTTC 93442
Qy 442 LysLeuTrpAlaPheGlyGlyIleMetPheGln----- 452
Db 93443 AAGCTATGGGCTTTCTTGGGATTTATGTTTCAGGTTAAAAAATTTACTAAACTGCTGCAGT 93502
Qy 453 -----ValPro 454
Db 93503 CGATTTTTTACTAAACTCTAATCTCATATCTTCATTCGACCAACCAATTTGTTTGAAGTGCCT 93562
Qy 455 LeuValLeuIleThrAsnTyrLeuGlnAsnLysPhe-ArgAsnSerMet----- 470
Db 93563 TTGTCTTCATCAAACTATCTACAGAAAGGTTTGGTCAACGGTATGCTCTCAAAAC 93622
Qy 470 ----- 470
Db 93623 CCGAGAAAATAGAAGAAATAACTCTTCTTTTCATAGCTAGCCATTTAAATCGCAATGCT 93682
Qy 471 -----ValGlyAsnM 474
Db 93683 GAAATTTAAATAAAGGTGATCTGTTTTTGAATGGGATCATATATTATAGTGGGGAACA 93742
Qy 474 etIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTrp 494
|||||

Db 93743 TGATCTCTGGTTCATCTTCGATTTTCGACACCGCATGTGTGTCTTCTTTATTACC 93802
QY 494 isAspLeuMetAsnArgLysGlyLysLeu 503
|||||
Db 93803 ACGACCTGATGAACCGAAGGATCGATG 93831
RESULT 15
AF078752 1650 bp mRNA linear ROD 11-NOV-1998
DEFINITION Mus musculus diacylglycerol acyltransferase (Dgat) mRNA, complete
cds.
ACCESSION AF078752
VERSION AF078752.1 GI:3859933
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1650)
AUTHORS Cases, S., Smith, S.J., Zheng, Y.-W., Myers, H.M., Lear, S.R., Sande, E.,
Novak, S., Collins, C., Welch, C.B., Lusis, A.J., Erickson, S.K. and
Farese, R.V. Jr.
TITLE Identification of a gene encoding an acyl CoA:diacylglycerol
acyltransferase, a key enzyme in triacylglycerol synthesis
Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13018-13023 (1998)
PUBMED 99007259
JOURNAL 9789033
MEDLINE 2 (bases 1 to 1650)
AUTHORS Farese, R.V. Jr.
TITLES Direct Submission
JOURNAL Submitted (16-JUL-1998) Gladstone Institute, 2550 23rd Street, San
Francisco, CA 94110, USA
FEATURES
source
1..1650
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="15"
1..1650
/gene="Dgat"
96..1592
/gene="Dgat"
/note="Dgat"
/codon_start=1
/product="diacylglycerol acyltransferase"
/protein_id="AAC72917.1"
/db_xref="GI:3859933"
/translation="MGDRGGAGSRRTSRVSGGGPKVDEDEVRDAVSPDLG
AGGAPAPAPAPAHTRDKGRTSVGDYVWLRCHPLQDLSFSSDSFGNRYGILNWCY
VMLILSNARPLENLIKYLVDPIQVYSLFLKDPYSWPAFCVIIASNIFVVAFOIE
KRLAVGALTEQMLLHVNLATIIICFFAAVALLVESIPDNLTLYRDIYFIFAPTLCEYL
SYRDVNLRCRRVRKAVSTGKRVSGAAQAVSYPDNLTYRDIYFIFAPTLCEYL
NPPSPRIKRFLLRRVLEMLFFQLOVGLIQQMVPYIQLNSMFKPKDMYSRIERL
LKLAVNLHLIWLIFWYWFHSCINAVAELOQDFREYRDMWNAESVYFWQNWNPV
HKWCFRHYKPLRHGSSKWKARTGVLTSAFFHEYLVSVDPLRMFLAFTAMMAQVP
LAWIVREFQNGYCNAAVWVLLIQPVAVLMVYHDYVLYNDAFVGV"
BASE COUNT 333 a 427 c 471 g 419 t
ORIGIN
Alignment Scores:
Pred. No.: 4,46e-62 Length: 1650
Score: 807.50 Matches: 190
Percent Similarity: 55.36% Conservative: 89
Best Local Similarity: 37.70% Mismatches: 174
Query Match: 30.13% Indels: 51
DB: 10 Gaps: 16
US-09-856-018b-16 (1-504) x AF078752 (1-1650)
QY 17 SerSerLeuArgArgPro----SerAlaThrSerThrAlaGlyLeuPheAsnSerPro 35
|||||
Db 120 AGCTCTCGCGCTCGGAGACCGCGCTCGGGGTTTCGTCAGGTT-----164

QY 36 GluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGlySerAspSer 55
|||||
Db 165 -----GGTAGTGGCCCAAGTAGAGAGGACGAGGTGGAGACGGCT 209
QY 56 IleAsnSerAsp-----AspAlaAlaValAsnSerGlnGlnAsnGlu 70
|||||
Db 210 GTGAGCCCGGACTTGGGCGCGGGGTGACGCGCGGCTCCGGCTCCAGCCCAT 269
QY 71 LysGlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerVal----- 87
|||||
Db 270 ACCCGGACAAAGAC-----GGCGGACCAAGCGCTGGCGGACGGC 308
QY 88 -----ProAlaHisArgLysValLysGluSerProLeuSerSerAspPhe 103
|||||
Db 309 TACTGGGATCTGAGTGCCTCGT-----CTGCAAGATCTTGTTCAGCTCAGACAGTGGT 365
QY 104 PheArgGlnSerHisAlaGlyLeuPheAsnLeuLysValLeuValAlaValAsn 123
|||||
Db 366 TTC---AGCAATATCTCGTATCTGATCTGATTTGGTGTGTGATGCTGATCTGAGTAAT 422
QY 124 SerArgLeuIleGluAsnLeuMetLysTyrGlyTrpLeuLysSerGlyPheTyr 143
|||||
Db 423 GCAAGGTTATTTTAGAGAAACCTTATCAAGTATGGCATCTCTGGTG---GATCCTATCCAG 479
QY 144 PheSerSerLysSerLeuArgasp-----TrpProLeuPheMetCysCysLeuSer 160
|||||
Db 480 GTGGTGTCTCTGTTTGAAGGACCCCTACAGCTGCCTGCCCATCGCTGATATTGCA 539
QY 161 LeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180
|||||
Db 540 TCCAATATTTTGTGTGCTGCTGATTTTCAGATTGAGAGCGCTGGCAGTGGTGGCTCCG 599
QY 181 ProGluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTyrPro 200
|||||
Db 600 ACAGAGCAGATGGGCTGCTGTACATGTGGTTAACTGGCCACATCATTTGCTCCCA 659
QY 201 ValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPhe 220
|||||
Db 660 GCAGCTGTGGCTTACTGTTGAGTCTATCACTCCAGTGGGTCTCGTGTGCTGGCA 719
QY 221 Ser---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyr----- 236
|||||
Db 720 TCATACTCCATCATGTCTCTCAAGCTTTATCTCAGGGAGTGTCAACCTGTGGTCCCG 779
QY 237 AspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuAspThrLeu--- 255
|||||
Db 780 CAGCGAAGGCTCAAGCCCAAGCTGCTCTACAGGAAGAAGGTCACTAGTGGGCTGTGCC 839
QY 256 -----AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVal 273
|||||
Db 840 CAGCAAGCTGTGAGCTATCCAGACACCTGACCTACCGAGATCTCTATTACTTCTT 899
QY 274 AlaProThrLeuCystyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTyr 293
|||||
Db 900 GCTCTACTTTGTGTATGAACCTCACTTCTCGGTGCCCGGAGATACAGAACGCCCTTT 959
QY 294 LeuPheArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheLeuLeuAsp 313
|||||
Db 960 CTGCTACGACGAGTCTCTGAGATGCTCTTTTACCACAGCTTCAAGTGGGCTGATCCAA 1019
QY 314 GlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuTyr 333
|||||
Db 1020 CAGTGGATGTCCTACTATCCAGAACTCCATGAAACCCCTTCAAG---GATATGGACTAT 1076
QY 334 Ala-----ThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCys 351
|||||
Db 1077 TCACGAGATCATGAGCGTCTCTTAAAGCTGGCGGTCCCAACCATCTGATCTGGCTATC 1136
QY 352 MetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuArgPheGly 371
|||||
Db 1137 TTCTTCTATTGTTTTCCTCCTCTCTCAATGCTGTGGCAGAGCTTCTGCGAGTTTGA 1196
QY 372 AspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMet 391

```
|||||
Db 1197 GACCGCGAGTTCTACAGAGATTGGTGGAACTGCTGAGTCTGACCTACTTTTGGCAGAAC 1256
Qy 392 TrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHis 411
|||||
Db 1257 TGGAAATATCCCGCGCACAACTGGTGCATCAGACACTCTACAAAGCCCTATGCTCAGACAT 1316
Qy 412 GlyLeuProLysAlaAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGlu 431
|||||
Db 1317 GGCAGCAGCAAAATGGGTGGCCAGGACAGGAGTATTTTGGCTCAGCCTTCTTCATGAG 1376
Qy 432 LeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPhe 451
|||||
Db 1377 TACCTAGTGAGCGTTCGCCCTCGCGATGTTCCGCTCTGGGCATTTCACAGCCCATGATGGCT 1436
Qy 452 GlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetVal 471
|||||
Db 1437 CAGGTCCCACTGGCC-----TGGATTGTGGCCGATTCCTCCAAGGGAACATAT 1484
Qy 472 GlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeu 491
|||||
Db 1485 GGCAATGCAGCTGTGTGG--GTGACACTCATCATTTGGGCAACCGGTGGTGTGTCATG 1541
Qy 492 TyrTyrHisasp 495
|||||
Db 1542 TATGTCCACGAC 1553
```

Search completed: February 21, 2003, 20:49:52
Job time : 3493 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 15:29:03 ; Search time 430 Seconds
(without alignments)
10170.653 Million cell updates/sec

Title: US-09-856-018b-15
Perfect score: 142
Sequence: 1 tagaaacacgctcggtctt.....aaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*						
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*						
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*						
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*						
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*						
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*						
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*						
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*						
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*						
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*						
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*						
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*						
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*						
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*						
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*						
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*						
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*						
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*						
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*						
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*						
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*						
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*						
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*						
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1942	100.0	1942	21 AAA48939	Soybean diacylglyc
2	719.8	37.1	1904	21 AAA51482	A. thaliana diacyl
3	719.8	37.1	1942	21 AAA88835	Arabidopsis acyl C
4	719.8	37.1	1942	21 AAZ45371	Acyl-CoA:cholester
5	719.8	37.1	1942	22 AAS01106	Arabidopsis thalia
6	716.6	36.9	1888	21 AAA48932	Arabidopsis diacyl
7	683	35.2	1985	21 AAA51484	A. thaliana AS11 d
8	672.8	34.6	1975	21 AAA48942	Wheat diacylglycer
9	630.2	32.5	1587	21 AAA48938	Rice diacylglycer

10	478.8	24.7	1281	21 AAA48933	Corn diacylglycer
11	413.4	21.3	901	21 AAA48936	Corn diacylglycer
12	404.8	20.8	1559	21 AAA48935	Corn diacylglycer
13	259.6	13.4	470	21 AAA48940	Soybean diacylglyc
14	253.2	13.0	380	24 ABQ85322	Arabidopsis thalia
15	235.2	12.1	978	21 AAA48934	Corn diacylglycer
16	229.8	11.8	629	21 AAA51486	EST with homology
17	229.8	11.8	629	21 AAZ45453	A. thaliana diacyl
18	199.4	10.3	234	21 AAA88839	Soybean acyl CoA:c
19	199.4	10.3	234	21 AAZ45372	Acyl-CoA:cholester
20	199.4	10.3	234	22 AAS01311	Soybean sterol acy
21	194.6	10.0	275	21 AAA88840	Soybean acyl CoA:c
22	194.6	10.0	275	21 AAZ45374	Acyl-CoA:cholester
23	194.6	10.0	275	22 AAS01313	Soybean sterol acy
24	187.2	9.6	267	21 AAA88841	Soybean acyl CoA:c
25	187.2	9.6	267	21 AAZ45373	Acyl-CoA:cholester
26	187.2	9.6	267	22 AAS01312	Soybean sterol acy
27	170.2	8.8	1650	21 AAZ49452	Mouse Diacylglycer
28	162.6	8.4	1521	19 AAU01533	Human acylcoenzyme
29	162.6	8.4	1976	21 AAU76169	Human ACAT Related
30	159.8	8.2	1766	21 AAA88846	Rat acyl CoA:chole
31	159.8	8.2	1766	21 AAZ45385	Acyl-CoA:cholester
32	159.8	8.2	1766	22 AAS01105	Rat sterol acyltra
33	159.4	8.2	1895	21 AAA88842	Human acyl CoA:cho
34	159.4	8.2	1895	21 AAZ45383	DNA encoding a pro
35	154	7.9	993	22 ABA06413	Human CDNA SQ ID
36	149.6	7.7	470	24 ABL93444	Arabidopsis thalia
37	140.6	7.2	983	19 AAU01539	Human acylcoenzyme
38	130.4	6.7	254	21 AAA48937	Rice diacylglycer
39	125.2	6.4	452	24 ABL81686	Human ovarian canc
40	124.4	6.4	253	21 AAA88849	Maize acyl CoA:cho
41	124.4	6.4	253	21 AAZ45376	Acyl-CoA:cholester
42	124.4	6.4	253	22 AAS01315	Maize sterol acylt
43	122.4	6.3	3996	23 ABL28621	Drosophila melanog
44	121.6	6.3	646	21 AAA48941	Wheat diacylglycer
45	121.2	6.2	1122	20 AAU30335	DNA encoding a hum

ALIGNMENTS

RESULT 1			
AAA48939	Standard	cdna	1942 BP.
ID	AAA48939 standard	cdna	1942 BP.
XX	AAA48939		
AC			
XX			
DT	06-DEC-2000 (first entry)		
DE	Soybean diacylglycerol acyltransferase cDNA #1.		
KW	Diacylglycerol acyltransferase; corn; rice; soybean; wheat;		
KW	triacylglycerol; herbicide; EC2.3.1.20; ss.		
XX			
OS	Glycine max.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	29..1543	
FT		/*tag=	
FT		/product=	Diacylglycerol_acyltransferase
XX			
PN	WO200032756-A2.		
XX			
PD	08-JUN-2000.		
XX			
PF	01-DEC-1999;	99WO-US28354.	
XX			
PR	02-DEC-1998;	98US-0110602.	
XX			
XX	31-MAR-1999;	99US-0127111.	
PA	(DUPO.) DU PONT DE NEMOURS & CO E. I.		
XX			
PI	Cahoon EB, Kinney AJ, Cahoon RE;		

XX WPI; 2000-412308/35.
DR P-PSDB; AAY94519.
XX
XX
PT Polynucleotides encoding diacylglycerol acetyltransferase, useful for
PT synthesis of triacylglycerols and increasing the level of oils in plant
XX seeds -
XX
PS Claim 4; Page 51; 62pp; English.
XX
CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,
CC soybean and wheat were screened for sequences with homology to a
CC putative acyl CoA cholesterol acyltransferase related gene from
CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
CC sapiens and Mus musculus. The cDNA clones identified from this process
CC were used to form complete diacylglycerol acyltransferase cDNA
CC sequences. The present sequence is soybean diacylglycerol
CC acyltransferase cDNA from clone srl.pk0098.a8. Diacylglycerol
CC acyltransferases are involved in the synthesis of triacylglycerols.
CC Alteration of the expression of the diacylglycerol acyltransferase
CC DNA can be useful for increasing the level of oils in plant seeds.
CC Inhibitors of diacylglycerol acyltransferase may be useful as
CC herbicides.
XX
SQ Sequence 1942 BP; 517 A; 425 C; 377 G; 623 T; 0 other;

Query Match 100.0%; Score 1942; DB 21; Length 1942;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAGAACAACACGCTGGTCTTCTCCAAATGGCGATTTCCGATGAGCCTGAAAGTGATGC 60
Db 1 TAGAACAACACGCTGGTCTTCTCCAAATGGCGATTTCCGATGAGCCTGAAAGTGATGC 60

Qy 61 CACTGCTCTCAACACACTTCCCTCGCGCGCGCTCCCTCCGCGCACCTCCACGCGGGCCT 120
Db 61 CACTGCTCTCAACACACTTCCCTCGCGCGCGCTCCCTCCGCGCACCTCCACGCGGGCCT 120

Qy 121 CTTCAATTCGGCTGAGACAACACCGACAGTTCGGGTGATGACTTGGCCAAAGGATTCGG 180
Db 121 CTTCAATTCGGCTGAGACAACACCGACAGTTCGGGTGATGACTTGGCCAAAGGATTCGG 180

Qy 181 TTCGACGACTCCATCAACAGCGAGGAGCGCGCGCTGCATTTCCCAACAGCAAAACGAAAA 240
Db 181 TTCGACGACTCCATCAACAGCGAGGAGCGCGCGCTGCATTTCCCAACAGCAAAACGAAAA 240

Qy 241 ACAAGACACTGATTTCCGCTCAAAATTCGCTACCGTCCCTCCGTCACCGCTCACCG 300
Db 241 ACAAGACACTGATTTCCGCTCAAAATTCGCTACCGTCCCTCCGTCACCGCTCACCG 300

Qy 301 CAAAGTGAAGAAAGTCGGCTCAGCTCCGACACTATTTTCGGTCAGAGTCACGCGGGCCT 360
Db 301 CAAAGTGAAGAAAGTCGGCTCAGCTCCGACACTATTTTCGGTCAGAGTCACGCGGGCCT 360

Qy 361 CTTCAACCTTTGATAGTAGTTCCTTGTGCTGAATAGCGGACATCATTTGAGAATTT 420
Db 361 CTTCAACCTTTGATAGTAGTTCCTTGTGCTGAATAGCGGACATCATTTGAGAATTT 420

Qy 421 AATCAAGTATGTTGGTTGATCAAAATCGGCTTTTGGTTTGGTTTCAAGTTCATTGAGAGA 480
Db 421 AATCAAGTATGTTGGTTGATCAAAATCGGCTTTTGGTTTGGTTTCAAGTTCATTGAGAGA 480

Qy 481 CTGCCCCCTTTTCATGTGTGCTTCTCTGTGGTATTTCTCTTCGGTGCCTTTATAGT 540
Db 481 CTGCCCCCTTTTCATGTGTGCTTCTCTGTGGTATTTCTCTTCGGTGCCTTTATAGT 540

Qy 541 GGAGAAGTGGCAACAGGAAGTGTATACCGGAACCGAGTTGTTGTTGTTGTTGTTGTTGTT 600
Db 541 GGAGAAGTGGCAACAGGAAGTGTATACCGGAACCGAGTTGTTGTTGTTGTTGTTGTTGTT 600

Qy 601 CATTAACCTCAACTTCGCTTTTCTATCAGATTTTATTTCTCAGGCTGATTCCTGCTTT 660
Db 601 CATTAACCTCAACTTCGCTTTTCTATCAGATTTTATTTCTCAGGCTGATTCCTGCTTT 660

Qy 661 TGTATCAGGTGTACAGTTAAATGCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
Db 661 TGTATCAGGTGTACAGTTAAATGCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720

Qy 721 TGCACATACAAACATATGATATGAGAGCACTTACCACAAATAGTTGAAAAGGAGAGCACT 780
Db 721 TGCACATACAAACATATGATATGAGAGCACTTACCACAAATAGTTGAAAAGGAGAGCACT 780

Qy 781 CTTGATGACTCTGAACATGGAGTATCTTACAACTAGCTTCAAGAGCTTGGCATATTT 840
Db 781 CTTGATGACTCTGAACATGGAGTATCTTACAACTAGCTTCAAGAGCTTGGCATATTT 840

Qy 841 CTTGTTGCCCTTACATATGATGTTACCAAGCTATCTCGCACACCTTATATTCGAAA 900
Db 841 CTTGTTGCCCTTACATATGATGTTACCAAGCTATCTCGCACACCTTATATTCGAAA 900

Qy 901 GGGTGGTGTGTTTCCGCCAATGTGCAAGCTGATATATTTACAGAGTATTTGGATTTAT 960
Db 901 GGGTGGTGTGTTTCCGCCAATGTGCAAGCTGATATATTTACAGAGTATTTGGATTTAT 960

Qy 961 AATAGACCAATATATTAATCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020
Db 961 AATAGACCAATATATTAATCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020

Qy 1021 TCTTTAGCCACCGAGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTGTGGCTCTG 1080
Db 1021 TCTTTAGCCACCGAGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTGTGGCTCTG 1080

Qy 1081 CATGTTCTATGCTTTTCCACCTTGGTTAAATATCTCGCAGAGCTTCTTCGATTTGG 1140
Db 1081 CATGTTCTATGCTTTTCCACCTTGGTTAAATATCTCGCAGAGCTTCTTCGATTTGG 1140

Qy 1141 TGATCGTGAATCTACAGAGATTTGGTGAATGCCAAAATCTGCGAAGATTTATTTGGAGAT 1200
Db 1141 TGATCGTGAATCTACAGAGATTTGGTGAATGCCAAAATCTGCGAAGATTTATTTGGAGAT 1200

Qy 1201 GTGGAATGATGCTGTTCACAAATGATGATCGGCACCTATATTTTCATGTTTAAAGCA 1260
Db 1201 GTGGAATGATGCTGTTCACAAATGATGATCGGCACCTATATTTTCATGTTTAAAGCA 1260

Qy 1261 CGGTCTACAAAGGCTGCTGCTCTTTAAATGCTTCTGCTTTCTGCTTTATTTCCATGA 1320
Db 1261 CGGTCTACAAAGGCTGCTGCTCTTTAAATGCTTCTGCTTTCTGCTTTATTTCCATGA 1320

Qy 1321 GCTGTGATTCCTGCTTCCGACATATTCAGTGTGGGCTTTCCGTTGAATATGTT 1380
Db 1321 GCTGTGATTCCTGCTTCCGACATATTCAGTGTGGGCTTTCCGTTGAATATGTT 1380

Qy 1381 TCAGGTTCCCTTTGGTCTTGTGATCACTAATTTCTGCAAAATAAATTCAGAAAATCAATGGT 1440
Db 1381 TCAGGTTCCCTTTGGTCTTGTGATCACTAATTTCTGCAAAATAAATTCAGAAAATCAATGGT 1440

Qy 1441 TGGAAATATGATTTTTTGGTTCATATTCAGTATCCTTGGTCAACCTATGTTGTTGCTGCT 1500
Db 1441 TGGAAATATGATTTTTTGGTTCATATTCAGTATCCTTGGTCAACCTATGTTGTTGCTGCT 1500

Qy 1501 ATACTACCATGACTTGATGAATAGAAAGCAAACTTGACTGAAGCTACGCCCATATACAT 1560
Db 1501 ATACTACCATGACTTGATGAATAGAAAGCAAACTTGACTGAAGCTACGCCCATATACAT 1560

Qy 1561 TTTAAAGTGTCACATGGATGAGCTTTTTCAGTTTTTCAGATTTGAAAATTTGATGTTGGATG 1620
Db 1561 TTTAAAGTGTCACATGGATGAGCTTTTTCAGTTTTTCAGATTTGAAAATTTGATGTTGGATG 1620

Qy 1621 TTGTTCAATATTTGTTTTCAGAAATGCTTTTCATCTACCAATGGCATTTGGCTGCTCTGAAG 1680
Db 1621 TTGTTCAATATTTGTTTTCAGAAATGCTTTTCATCTACCAATGGCATTTGGCTGCTCTGAAG 1680

Qy 1681 GAATTCACCGGATATGCCAGTTCACAGGCTAATTCATATCTTGATCTATGTTACTTAC 1740
Db 1681 GAATTCACCGGATATGCCAGTTCACAGGCTAATTCATATCTTGATCTATGTTACTTAC 1740

Db 835 TTTATCAGGTGTCACTTTGATGCTCCTCCTCACTTGTCATTTGTGGCTAAAGTTGGTTTCCTTA 894
Qy 721 TGCACATACAAACTATGATATGAGACACTTACCAAAATAGTTGAAAGGAGGAGCACT 780
Db 895 TGCTCATACTAGCTATGATCAAGATCCTAGCCAACTGACCTGATAGG----- 944
Qy 781 GCTCGATACTCTGAACATGGACTATCCTTACCACTAAGCTTCAAGAGCTGGCAATATT 840
Db 945 -----CCAATCTGAAGTCTCTCTACTACGTTAGCTTGAAGAGCTTGGCATATT 993
Qy 841 CCTGTTCCCTCCCTACATATGTTACCAGCAAGCTATCTCCACACCTTATATTCGAAA 900
Db 994 CATGTCCTCCACATATGTTATACCAAGTTATCCAGCTTCTGTCATGATATACGAAA 1053
Qy 901 GGGTGGTGTGTTCCCACTTGTCAAGCTGATATATTTACAGAGTTATGGGATTTAT 960
Db 1054 GGGTGGTGTGCTGTCATTTGCAAACTGTCATATTCACCGGATTCATGGGATTTAT 1113
Qy 961 AATGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020
Db 1114 AATGACCAATATATAAATCCTATTTGTGAGGAACCTCAAAAGCATCTTTGAAAGCGCATCT 1173
Qy 1021 TCTTTACCCACCGAGAGAGTCTCAAGCTTCTGTTCCAAATTTATATGTTGGCTCTG 1080
Db 1174 TCTATATCTATTTGAAGAGTGTGAAGCTTTCAGTTCCAAATTTATATGTTGGCTCTG 1233
Qy 1081 CATGTTCTATGCTTTTCCACCTTGGTTAAATATCTGTCAGAGCTTCTTCGATTTGG 1140
Db 1234 CATGTTCTACTGCTTCTCCACCTTGGTTAAACATATTCGACAGCTTCTGCTCGG 1293
Qy 1141 TGATCGTAATCTACAAAGGATTTGGTGAATGCGCAAACTGTCGAAGATTTATGGAGAT 1200
Db 1294 GGATCGTGAATCTACAAAGATTTGGTGAATGCGCAAACTGTCGAAGATTTATGGAGAT 1353
Qy 1201 GTGAATATGCTGTTCACAAATGATGATCGGCACCTATATTTCCATGTTTAAAGCA 1260
Db 1354 GTGAATATGCTGTTCACAAATGATGATGCTGACATATATCTCCGCTGCTGCGAG 1413
Qy 1261 CGGTCTACCAAGGCTGCTCTTTAAATGCTTCTGTTTCTGCTTTTATTCATGA 1320
Db 1414 CAAGATACCAAGACACTCGCCATATCATTTGCTTCTAGTCTGTCAGTCTTTCATGA 1473
Qy 1321 GCTGTGATGCTGCTTCTGTCACATATTCAGTTGTGGCTTTCGTTGGAATATGTT 1380
Db 1474 GCTATGATCGAGTCTCTGCTCTCTCAAGCTATGCGCTTTTCTGTTGGATATGTT 1533
Qy 1381 TCAGTTCCTTTGCTTGTGATCACTAATATCTGCAAAATAAATTCAGAACTCAATGTT 1440
Db 1534 TCAGTTCCTTTGCTTGTGATCACTAATATCTGCAAAATAAATTCAGAACTCAATGTT 1590
Qy 1441 TGAATAATGATTTTGGTTTCATATTCAGTATCTTGGTCAACCTATGTTGCTGCT 1500
Db 1591 GGGAAACATGATCTCTGTTTCATCTCTGTCATTTTCGACAAACCGATGTTGCTGCTCT 1650
Qy 1501 ATACTACCATGATGATGATAGAAAGGAACTTGCAGTGAAGCTACGG 1551
Db 1651 TTATTAACACAGCTGATGAACCGAAAGGATGATGTCATGTAACAACTG 1701

RESULT 7

AA51484

ID AA51484 standard; cDNA; 1985 BP.

XX

AC AA51484;

XX

DT 09-OCT-2000 (first entry)

XX

DE A. thaliana AS11 diacylglycerol acyltransferase cDNA.

XX

KW DGAT; diacylglycerol acyltransferase; seed oil; fatty acid synthesis;

KW size; weight; carbon flux; TAG1; insertion mutant; ss.

XX

OS Arabidopsis thaliana.

XX

FH Key

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

XX

DN WO200036114-A1.

XX

XX

PD 22-JUN-2000.

XX

XX

PF 16-DEC-1999; 99WO-CA01202.

XX

PR 17-DEC-1998; 98US-0112812.

XX

PA (CANA) NAT RES COUNCIL CANADA.

XX

PI Zou J, Taylor DC, Wei Y, Jako CC;

XX

XX

DR WPI; 2000-431592/37.

XX

XX

P-PSDB; AAY96854.

XX

PT New DNA encoding diacylglycerol acyltransferase from Arabidopsis

PT

PT

PT

PT

PT

XX

PS Claim 5; Page 79; 91pp; English.

XX

XX

CC This cDNA encodes Arabidopsis thaliana ecotype Columbia mutant AS11

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ Sequence 1985 BP; 482 A; 415 C; 446 G; 642 T; 0 other;

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

35.2%; Score 683; DB 21; Length 1985;

74.4%; Pred. No. 1.7e-133;

0; Mismatches 285; Indels 24; Gaps 2;

Qy 345 AGAGTCAGCGCGGCTCTTCAACCTTTGTATAGTAGTCTTGTGCTGTAATAGCCGAC 404

Db 608 AGAGCCATGCCGGATTTCAACCTCTGTGTAGTAGTTCTTATTGCTGTAACAGTAGAC 667

Qy 405 TCATCATTTGAGAATTTAATGAAGTATGTTGTTGATCAAAATCTGGCTTTTGGTTTACTT 464

Db 668 TCATCATCGAAAATCTTATGAAGTATGTTGTTGATCAGAACGGATTTCTGGTTAGTT 727

Qy 465 CAAAGTCATGAGAGACTGGCCCTTTTCATGTTGTTGCTCTCTCTGTTGTTATTCCTT 524

Db 728 CAAGATCGTCGAGATTTGGCGCTTTTCATGTTGTTGATATCCCTTTTCGATCTTTCCCT 787

Qy 525 TCCTGCTTTTATAGTGGAGAAGTTGGCACAACGGAAGTGTATACCCGAAACCTGTTG 584

Db 788 TGGCTGCTTTTACGGTTGAGAAATTTGGTACTTCAGAAATACATATCAGAACCTGTTGCA 847

Qy 585 TTGTACTTCATATATATCATTTACCTCAACTTCCTTTTCTATCCAGTTTACTTATTCCTCA 644

Db 848 TCTTCTTCATATATATATCACCATGACAGAGTTTGTATCCAGTTTACCTACCCCTAA 907

Qy 645 GGTGTGATTCGCTTTTGTATATCAGGTGTCAGGTTAAATGCTGTTTCTTGTGTTGATGGT 704

```
Db 908 GGTGTGATTCGCTTTTATCAGGTGTCACCTTGTATGCTCCTCAGCTTGCATTTGTGGC 967
Qy 705 TAAATTTGGTCTTATGACATCAAACTATGATATGACGACCACTTACCAAAATAGTTG 764
Db 968 TAAAGTTGGTTCTTATGCTCATACTAGTATGACATAAGATCCCTAGCCAACTGACGCTG 1027
Qy 765 AAAAGGNGAAGCACTGCTCATCTCTGAACATGGAGCTATCCTTACAACTAAGCTTCA 824
Db 1028 ATAAAG-----CCAACTCCTGAAGTCTCCTACTACCTTACCTAGCTTGA 1066
Qy 825 AGAGCTGGCATATTTCTGTGTCGCCCTACATATGTTACCAAGCAGCTATCCTCGCA 884
Db 1067 AGAGCTGGCATATTTCTGTGTCGCCCTCCTCACTTGTGTTATCAGCCAGTATCCACGTT 1126
Qy 885 CACCTTATATTCGAAAGGTTGGTTGTTCCGCAACTTGTCAAGCTGATATATTTACAG 944
Db 1127 CTGATGATATACGGAAGGTTGGTGGCTCGTCAATTTGCAAACTGCTCATATTCACCG 1186
Qy 945 GAGTTATGGGATTTATATAGACCAATATATTAATCCCATAGTACAAATTCACAGCATC 1004
Db 1187 GATTCATGGGATTTATATAGAACATATATAAATCCCTATTTGTCAGGAACCTCAAGCATC 1246
Qy 1005 CTCTCAGGGAAACCTTCTTTACGCCACCGAGAGAGTCTCAAGCTTCTCTTCCAAAT 1064
Db 1247 CTTTGAAGGCGATCTCTATATGCTATGTAAGAGAGTGTGAAGCTTTCAGTTCCAAAT 1306
Qy 1065 TATATGTGCTGCTGCTATGCTTCTATGCTTTTCCACCTTTGTTAAATATCCTGGCAG 1124
Db 1307 TATATGTGCTGCTGCTATGCTTCTATGCTTCTTCCACCTTTGTTAAATATGCGAG 1366
Qy 1125 AGCTTCTCGATTTGGTATCGTGAATCTACAGGATTTGTTGAATGCCAAAAGTCTG 1184
Db 1367 AGCTTCTGCTGCTGGGATCGTGAATCTACAAAGATTTGTTGAATGCCAAAAGTCTG 1426
Qy 1185 AAGATTATGAGATGTTGGAATATGCTGTTCAAAAGTATGTTGAATGCCAAAAGTCTG 1244
Db 1427 GAGATTCTGGAGATTTGGAATATGCTGTTCAAAAGTATGTTGAATGCCAAAAGTCTG 1486
Qy 1245 TTCCATGTTTAAAGCAGCTGTACAAAGGCTGCTCTTTTAAATGCTTCTCTGTTT 1304
Db 1487 TCCGCTGCTTCCGCGCAGCAAGATACCAAGACACTCGCCATTTATCATTTCTCTAGTCT 1546
Qy 1305 CTGCTTATTCAGCTGTGATGCTGTTTCTTCCACATATTTCAAGTTGTTGGGCTT 1364
Db 1547 CTGAGTCTTTTCATGAGCTATGCTATGCTGCTGCTTCTTCTTCAAGCTATGGGCTT 1606
Qy 1365 TCGTGGAAATATGTTTCTGCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1424
Db 1607 TCTTGGATTTATTTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1666
Qy 1425 TCAGAACTCAATGTTTGGAAATATGATTTTGTGTTTCTATATTCAGTATCCTTGTCAAC 1484
Db 1667 TTGG---CTCAAGCTGTTGGAATATGCTTCTGTTTCTGCTTCTGCTTCTGCTTCT 1723
Qy 1485 CTATGTTGTTACTGCTATATCTACCATGCTGATGATGATGATGATGATGATGATGAT 1544
Db 1724 CGATGTTGTTGCTTCTTATTTACCAAGCTGATGATGATGATGATGATGATGATGAT 1783
Qy 1545 GCTACGG 1551
Db 1784 ACACTG 1790
```

RESULT 8

AAA48942

ID AAA48942 standard; cDNA; 1975 BP.

XX

AC AAA48942;

XX

DT 06-DEC-2000 (first entry)

XX

DE Wheat diacylglycerol acyltransferase cDNA #2.

XX KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
KW triacylglycerol; herbicide; EC2.3.1.20; ss.
OS Triticum aestivum.
XX Key Location/Qualifiers
FT CDS 107..1633
FT /*tag= a
FT /product= Diacylglycerol_transferase
XX WO200032756-A2.
XX 08-JUN-2000.
XX 01-DEC-1999; 99WO-US28354.
XX 02-DEC-1998; 98US-0110602.
XX 31-MAR-1999; 99US-0127111.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Cahoon EB, Kinney AJ, Cahoon RE;
XX WPI; 2000-412308/35.
XX P-PSDB; AAY94522.
XX Polynucleotides encoding diacylglycerol acyltransferase, useful for
XX synthesis of triacylglycerols and increasing the level of oils in plant
XX seeds -
XX Claim 4; Page 56; 62pp; English.
XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,
XX soybean and wheat were screened for sequences with homology to a
XX putative acyl CoA cholesterol acyltransferase related gene from
XX Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
XX sapiens and Mus musculus. The cDNA clones identified from this process
XX were used to form complete diacylglycerol acyltransferase cDNA
XX sequences. The present sequence is wheat diacylglycerol
XX acyltransferase cDNA from clone wrl.p0119.b6.fis. Diacylglycerol
XX acyltransferases are involved in the synthesis of triacylglycerols.
XX Alteration of the expression of the diacylglycerol acyltransferase
XX DNA can be useful for increasing the level of oils in plant seeds.
XX Inhibitors of diacylglycerol acyltransferase may be useful as
XX herbicides.
XX Sequence 1975 BP; 463 A; 497 C; 461 G; 550 T; 4 other;

Query Match 34.6%; Score 672.8; DB 21; Length 1975;
Best Local Similarity 71.3%; Pred. No. 2.4e-131;
Matches 887; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

Qy 286 CGTCCCGCTCACCGCAAGTGAAGGAAAGTCCGCTCAGCTCCGACACTATTTCCGTC 345
Db 373 CGGCGCCGCCACCGCGGGTCAAGAGAGCGCCGTAGCTCCGACGCTTCCGACA 432
Qy 346 GAGTCACGCGGCTCTTCAACCTTTGTATAGTACTCCTTGTGCTGTAATAGCCGACT 405
Db 433 GAGCATGCAAGTCTCTGAACTATGATGTTGCTGCTGATGCACTGACGACGAGCT 492
Qy 406 CATCATGAGAAATTAATGAAAGTATGTTGTTGATCAAAATCTGGCTTTGGTTAGTTC 465
Db 493 CATTATCGAAGCTTAATGAAGTATGGCTTATTAAGAGCTGGGTTTGGTTAGTGC 552
Qy 466 AAGTCATTGAGACAGTGGCCCTTTTCATGTTGTTGCTTTCTCTTGTGGTATTTCCCTT 525
Db 553 AAGATCGCTGGGAGATTTGGCCACTTCTGATGCTGCTCTTACCATTTTCCCACT 612
Qy 526 CGCTGCCTTTATAGTGGAGAAAGTTGGCAACAAGGAGTATACCCGAACAGTTGTTGT 585
Db 613 TGCTGCTCTATGACCGAAGTGGGCTCAAGAAAGCTATCCGTGATCATGTGCTAT 672

Db	266	TTGCTTTAGTCTGCCCTGCTTTCCCCCTGGGTGCATTTGCGATTTGAAAGTTGGCAATTAA	325
Qy	559	GAAGTGTATACCCGAACCAAGTGTGTGTCTACTTATATATCATATTAACCTCAACTTCGCCT	618
Db	326	CAATGTTATTACTGATGCTGTGCTACCTGCCTCCATATCTTCTTTTCAACAACCGAAAT	385
Qy	619	TTTCTATCCAGTTTATAGTTATCTCCAGGTGTGATTTCTGCTTTTGTATATCAGGTGTACGTT	678
Db	386	TGTATATCCAGTGTGTGATTCCTTTAAAGTGTGATTTCTGCAGTTTGTCTGGCTTTTGTGT	445
Qy	679	AATGCTGTTTCTGTGTGTATGCTTTAAAAATGGTGCTTTATGCACATACAAACTATGA	738
Db	446	GATATTTATGGCTGTATTTGTTTGGCTGAAGCTTGATATCTTTGCAATACAAACCAATGA	505
Qy	739	TATGAGAGCACTTACCAAAATTAGTTTGAAGGGAGAACCACTGCTCGCATCTCTGAACAT	798
Db	506	TATAAGGCAACTGACCATGGCGGCAAGAGGTTGATAATCAACTAAGCAGCTTGACAT	565
Qy	799	GGACTATCTTACAAGGTAAAGCTTCAAGAGCTTGGCATATTTCTCTGGTTGCCCTACATT	858
Db	566	GGATAATTTACAACCTTCCAACCTTAAAGGAATCTAATACTTCAATGATGGCTCTACACT	625
Qy	859	ATGTTACCAGCCCAAGCTATCCTTCGCACACCTTATATTTCCGAAGGGTGGTTGTTCCGCA	918
Db	626	CTGTTATCAGCCCAAGCTATCCCCGAACCTCATGTGTTAGAAAGAGTGGCTGATTCGTCA	685
Qy	919	ACTTGTCAAGCTGATAAATATTTACAGGAGTATTGGATTTTATAATAGACCAATATATTA	978
Db	686	AAATTATCTGACTTGATCTTTTACTTGGTCTTCAAGGCTTCATTATTTGAGCAATACATAA	745
Qy	979	TCCCATAGTACAAAATTCACAGCATCCCTCTCAAGGGAAACCTCTTTTACGCCACCCGAG	1038
Db	746	TCCAATTTGTTGAATCTTCAGCATCCCAATTGAAGGAGGAGCTCCTCTAAATGCTGTAGAC	805
Qy	1039	AGTCTCTGAAGCTTCTGTGTTCCAAATTTATATGTGTGGCTCTGCGATGTTCTATGCTTTT	1098
Db	806	TGTTTTGAAACTCTCAATTACCAAAATGTTTACTCTGGCTTTGCGATGTTCTATGCTTTT	865
Qy	1099	CCACCTTTGGTTAAATATCCTCGCAGAGCTTCTTCGATTTGGTGTGATCGTGAATTCACAA	1158
Db	866	CCATCTCTGTTAAGTATACTTTGCTGAGATCTTCGATTTGGTGACCGTGAATTCACAA	925
Qy	1159	GGATTGGTGGAAATGCCAAAACCTGTCGAAGATTTATGGAGGATGTGAATATGCTGTTC	1218
Db	926	AGATTGGTGGAAATGCAAAAACAAATTTGATGAGTATTTGGAGAAATATGCGCTGTACA	985
Qy	1219	CAAAATGGATCCGCCACCTATATTTTCCATGTTTAAAGGCACCGTCTACCAAGGCTGC	1278
Db	986	TAAATGGGTGTTCGCCATATTTACTTTTTTCTTGGCATGCGAAATGGTATATCAAGGAAGT	1045
Qy	1279	TGCTCTTTTAAATGGCTTCTGTTTCTGCTTTATTTCCATGAGTGTGCGATTTGCTGTTC	1338
Db	1046	TGCTGCTTGATATCATCTCTTTGTTTCTGGCGTACTCCATCAGATATGTGTGCGCTGTTC	1105
Qy	1339	TTGCCATATATTCAGTTGTGGGCTTTCGGTGGAAATATGTTTTCAGGTTTCTTTGGTCTT	1398
Db	1106	CTCGCGCATTTCTCAAGTTCTGGGCATTTCTTTAGGAATTAATGCTACAGATCCCCCTTTCGT	1165
Qy	1399	GATCACTAATTATCTGCAAAATTAATTTAGAAACTCAATGTTTGGAAATATGATTTTTTG	1458
Db	1166	ATTGACAGCATCTCTCAAAAGTAAATTTCAGAGATACAAATGGTTGGCAACATGATATTTT	1225
Qy	1459	GTTTCATATTCAGTATCTTTGGTCAACCTATGTGTCTACTGCTATCTACATGACTGATTTG	1518
Db	1226	GTTCTTTTCTGCACTATTTGGCAGCGCAATGTGCCCTTCTCTCTGTACTATCATGATGTGAT	1285
Qy	1519	GAATAGGAAGGCAA	1533
Db	1286	GAACAGGATTGAGAA	1300
RESULT 10			
AAA48933			

Db 61 TTTCCCTTGGTGCATTTGGAGTCGAAAGTTGGCATTCACAAATCTCGTTAGTGATCCT 120
Qy 578 GTTGTGTGTACTATATATATACCAATTCAGTTCGCTTTTCTATCCAGTTTTAGTT 637
Db 121 GCTACTACCTGTTTTCACATCCTTTTACAAATTTGAAATTTGATATATCCAGTGTCTG 180
Qy 638 ATTCACAGTGTGATTCGCTTTTGTATACAGGTGTCAGGTTAAAGTCTGTTTCTGTTGTT 697
Db 181 ATTCCTTAAGTGTGATTCGCAATTTTATCAGGCTTTGTGTGATGTTTATTCGCTGCATT 240
Qy 698 GTATGTTAAATTTGGTCTTTATGCACATACAAATCATATATGA-CAGCAGCTTACCAA 756
Db 241 GTTGGCTGAAGCTGTATGTTTTCACATACAAACCATGATATGAAGAAACATGATCAC 300
Qy 757 ATTAGTTGAAAGGAGAGCAGCTGCTGATCTCTGAACATGGAATCTCTTACAAAGT 816
Db 301 AAGCGCAAGAGGTTGATATGAATGAACCTGACCGCGCTGGCATAGATATTTACAACTCC 360
Qy 817 AAGCTTCAAGAGCTTGGCATATTTCTCGTGGTGGCCCTACATTTATGTTACCAAGCAAGCTA 876
Db 361 AACTCTTGGAGTCTAACATATCTCAAGATGGCTCCGACACTCTGTTATCAAGCCAAAGT 420
Qy 877 T---CCTCGCACACTTATATTCGAAAGGTTGGTTTTCGCCAAGCTGTCAGCTGAT 933
Db 421 TATCTNCGAACACTTATGTTAGAAAGGTTGGTGGCTGCGTCAAGTTATTCATATCTT 480
Qy 934 AATATTTACAGAGGTTATGGATTTATATAGACCAATATATTAATCCATAGTACAAAA 993
Db 481 GATATTTACTGCTCCAAAGGATTCATTTATGAGCAATACATAATCTATTTGTTGAA 540
Qy 994 TTCACAGATCCTCTCAAGGAAACCTTCTTTACGCCACCGAGAGGTTCTGAAGCTTTC 1053
Db 541 CTCTCAACATCCATGATGGAGGATTTACTGAATGCTGTAGAGACTGTTTGAAGCTCTC 600
Qy 1054 TGTCCAAATTTATATGTTGCTGCTGATCTCTATGCTTTTCCACCTTTGGTTAAA 1113
Db 601 ATTACCAATGCTACCTGTGGCTTGCATGTTTATGCTTTTCCATCTGTGTTAAA 660
Qy 1114 TATCTGCGAGAGCTCTCTCGATTTGTTGATGCTGGAATTTACAGGATTTGGTGAATGC 1173
Db 661 CATCTTGTGAGATCTCTCGATTTGTTGACCGAGAAATTTACAAAGACTGGTGAATGC 720
Qy 1174 CAAACTGTCCGAATTTATGAGAGATGTTGAATATGCTCTTCCAAATGGATGATCCG 1233
Db 721 AAAGACAAATTTGATGAGTACTGGAGAAATGGAACATGCTGTGCATTAATGGATTTGCG 780
Qy 1234 CCACCTATATTTTCCATGTTTAAAGCGAGGCTTACCAAGGCTGCTGCTCTTTTAATGTC 1293
Db 781 TCATATATATTTCCCTGTCATGCGAATGGTATATCAAGGAAGTTGCTGTTTTATATC 840
Qy 1294 CTTCCTGTTTCTGCTTTATTTCCATGAGCTGTGATTTGCTTCTTCCGACATATTCAA 1353
Db 841 GTTCTTTGTTCTGCTGTACTGATGAGTTATGTTGTTGCTGCTTCCCTGCCACATCTCAA 900
Qy 1354 GTTGTGGCTTTCCGTTGGAATTTGTTTTCAGTTCCTTGGTCTGTGATCATCAATATCT 1413
Db 901 GTTGTGGCTTTCTAGGAATTCATGCTTCAGATTCCTCATCATATTTGATCATCATCT 960
Qy 1414 GCAAAATTAATTCAGAACTCAATGGTTGGAATATGATTTTGGTTTCATATTCAGTAT 1473
Db 961 CAAAATTAATTCAGTGCACAAATGGTTGGCAATATGATCTTTTGGTTTTTCTGAT 1020
Qy 1474 CCTTGGTCAACCTATGTTGTGCTATATCTACCTGACTTGTATGATAGGAAGGCAA 1533
Db 1021 ATACGGGCGCAATGTGTGTTCTATTGTTATTTACCATGATGTGATGAACCGGACTGAGAA 1080.

RESULT 11
AAA48936
ID AAA48936 standard; cdna; 901 BP.
XX
AC AAA48936;
XX

06-DEC-2000 (first entry)
Corn diacylglycerol acyltransferase cDNA #4.
Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
triacylglycerol; herbicide; EC2.3.1.20; ss.
Zea mays.
Key Location/Qualifiers
CDS 1..858
/tag= a
/partial
/transl_except= (pos:442..444,aa:Xaa)
/transl_except= (pos:490..492,aa:Xaa)
/transl_except= (pos:587..589,aa:Xaa)
/transl_except= (pos:805..807,aa:Xaa)
/transl_except= (pos:820..822,aa:Xaa)
/note= "Xaa= unknown"
/product= Partial_diacylglycerol_acyltransferase
WO200032756-A2.
08-JUN-2000.
01-DEC-1999; 99WO-US28354.
02-DEC-1998; 98US-0110602.
31-MAR-1999; 99US-0127111.
(DUPO) DU PONT DE NEMOURS & CO E I.
Cahoon EB, Kinney AJ, Cahoon RE;
WPI: 2000-412308/35.
P-PSDB: AAY94516.
Polynucleotides encoding diacylglycerol acetyltransferase, useful for
synthesis of triacylglycerols and increasing the level of oils in plant
seeds -
Claim 18; Page 46; 62pp; English.
In the present invention, cDNA libraries from Arabidopsis, corn, rice,
soybean and wheat were screened for sequences with homology to a
putative acyl CoA cholesterol acyltransferase related gene from
Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
sapiens and Mus musculus. The cDNA clones identified from this process
were used to form complete diacylglycerol acyltransferase cDNA
sequences. The present sequence is corn diacylglycerol acyltransferase
cDNA from a contig of clones p0042.cspaf49r, p0122.chkmb57r and
p0125.czbauf61rb. Diacylglycerol acyltransferases are involved in the
synthesis of triacylglycerols. Alteration of the expression of the
diacylglycerol acyltransferase DNA can be useful for increasing the
level of oils in plant seeds. Inhibitors of diacylglycerol
acyltransferase may be useful as herbicides.
Sequence 901 BP; 256 A; 172 C; 193 G; 276 T; 4 other;
Query Match 21.38; Score 413.4; DB 21; Length 901;
Best Local Similarity 70.1%; Pred. No. 3.9e-77;
Matches 568; Conservative 0; Mismatches 239; Indels 3; Gaps 1;
Qy 712 GGTGTCTTATCCATACAAAATGATGATGAGAGCACTTACCAAAATAGTTGAAAAGGG 771
Db 30 GGTCTCTTATCCATACAAAATGATGATGAGAGCACTTACCAAAATAGTTGAAAAGGG 89
Qy 772 AGAAGCACTGCTCGATCTCTGAACATGGAGCTATCTTACAACTGTAAGCTTCAAGAGCTT 831
Db 90 TGCTGCATATGGAATTTATGTCGATCTCTGAGATATGAAGATCCCACTTTAAAAGTCT 149
Qy 832 GGCATATTTCTGTTGGCCCTACATTTATGTTACCAGCAAGCTATCTCTCGCACCTTA 891

Db	150	AGTGACTTCATGTTGGCCCCAACACTTTGTGTACAGGCCAACTTATCTCAAACATACATG	209
Qy	892	TATTCGAAGGGTTGGTTGTTTCGCCAACTTCTCAAGCTGATAATATTATACAGGAGTTAT	951
Db	210	TATTAGAAGGGTTGGGTGACCCAGCAACTCATAAAGTGGTGGTTTTTACAGGCTTCAT	269
Qy	952	GGGATTTATAATAGACCAATATATTAAATCCCATAGTACAAAATTCACAGATCTCTCAAA	1011
Db	270	GGGCTTCATAAATTTGAGCAATATATAAACCAATCTGGAAGAAATTCCAAACATCCACATGAA	329
Qy	1012	GGGAACCTTCTTTTACGCCACCGAGAGAGTCTGAAGCTTTCTGTGTCCAAATTTATATGT	1071
Db	330	AGGGAATTTTTCGAATGCTATAGAAGAGCTCTTAAACTCTCAGTCCCAACATTTATATGT	389
Qy	1072	GTGGCTCTGCATGTTCTATATGCTTTTTCACCTTTTGGTTAAATATCTGGCAGAGCTTCT	1131
Db	390	ATGGCTTTGTCATGTTCTATTTGCTTTTTCATTTATATGCTGAACATTTGAGCTTAACTCTCT	449
Qy	1132	TCGATTTGGTGATCGTGAATCTCACAGGATTTGGTGAATGCCAAAAGCTGTCGAAGTTA	1191
Db	450	CTGTTTCGGTGACCGTGAATCTTAAGGACTGGTGGAAATGNCAAAATCTGTTGAAGAGTA	509
Qy	1192	TTGAGGATGTGGAATATGCTCTTCACAAATGGATGATCGGCCACCTATATATTTCCATG	1251
Db	510	CTGGAGGATGTGGAACATGCTGTTTCATAAGTGGATCATCAGACACATATATTTCCATG	569
Qy	1252	TTTAAGCAGCGTCTACAAAAGGCTGCTGCTCTTTTAATTTGGCTTCCTGGTTTCTCGCTTT	1311
Db	570	TATAAGNAAAGCTTTTTCAGGGGTGATAGCTATTCTAATCTCGTTTCTGGTTTCAGCTGT	629
Qy	1312	ATTCCATGAGCTGTGCATTTGCTTCTTTCGCCACATATTTCAAGTTGTGGCTTTTCGGTGG	1371
Db	630	ATTCCATGAGATATGATTTGGGTGCGGTGCCACATTTCAAATCTCGGGCATTTTCTGG	689
Qy	1372	AATATGTTTTCAGTTCCTTTGGCTCTTGATCACTAATTTATCTGCAAAATAAATTCAGAAA	1431
Db	690	GATCATGTTTTCAGATACCGTTGGTATTTCTTTGAAGAGATATCCATGTCTAGTTTCAAGCA	749
Qy	1432	CTCAATGGTTTGGAAATATGATTTTTTTGGTTTCATATTCAGTATCTTGTGTCAACCTATGTG	1491
Db	750	TGTAAATGGTGGGCAACATGATATTTTGGTTC---TTCAGTATAGTCCGACACCGCATGTN	806
Qy	1492	TGTACTGCTATACTACCATGACTTGATGAA	1521
Db	807	GTGTCTCTATAACTAACTAAGCTAGCTAATAA	836

RESULT 12

AAA48935	standard; cDNA; 1559 BP.
AAA48935;	
06-DEC-2000	(first entry)
Corn diacylglycerol acyltransferase cDNA #3.	
Diacylglycerol acyltransferase; corn; rice; soybean; wheat;	
triacylglycerol; herbicide; EC2.3.1.20; ss.	
Zea mays.	
Key	Location/Qualifiers
CDS	1..984
FT	/*tag= a
FT	/partial
FT	/product= Partial_diacylglycerol_acyltransferase
WO200032756-A2.	
08-JUN-2000.	
01-DEC-1999;	99WO-US28354.

XX	02-DEC-1998;	98US-0110602.	
XX	31-MAR-1999;	99US-0127111.	
XX	(DUPO) DU PONT DE NEMOURS & CO E I.		
XX	Cahoon EB, Kinney AJ, Cahoon RE;		
XX	WPI: 2000-412308/35.		
XX	P-PSDB; AAY94515.		
XX	Polynucleotides encoding diacylglycerol acetyltransferase, useful for		
XX	synthesis of triacylglycerols and increasing the level of oils in plant		
XX	seeds		
XX	Claim 4; Page 44-45; 62pp; English.		
XX	In the present invention, cDNA libraries from Arabidopsis, corn, rice,		
XX	soybean and wheat were screened for sequences with homology to a		
XX	putative acyl CoA cholesterol acyltransferase related gene from		
XX	Arabidopsis thaliana and diacylglycerol acyltransferases from Homo		
XX	sapiens and Mus musculus. The cDNA clones identified from this process		
XX	were used to form complete diacylglycerol acyltransferase cDNA		
XX	sequences. The present sequence is corn diacylglycerol acyltransferase		
XX	cDNA from clone cp1lc.pk005.h23. Diacylglycerol acyltransferases are		
XX	involved in the synthesis of triacylglycerols. Alteration of the		
XX	expression of the diacylglycerol acyltransferase DNA can be useful for		
XX	increasing the level of oils in plant seeds. Inhibitors of		
XX	diacylglycerol acyltransferase may be useful as herbicides.		
XX	Sequence 1559 BP; 413 A; 299 C; 288 G; 555 T; 4 other;		
XX	Query Match	20.8%;	Score 404.8; DB 21; Length 1559;
XX	Best Local Similarity	66.5%;	Pred. No. 2.8e-75;
XX	Matches 580; Conservative	0;	Mismatches 292; Indels 0; Gaps 0;

[illegible]

Qy 954 GATTTATTAATAGACCAATATATTAATCCCATAGTACAAAAATTCACAGCATCCTCTCAAGG 1013
Dy 70 GATTCATTAATTGACCAATACATAATCCTATTGTGTGAACCTCTCAACATCCATTGATGG 129
Qy 1014 GAAACCTTCTTTAGCCACCGAGAGAGTTCTGAAGCTTTCTGTCCAAATTTATATGTGT 1073
Dy 130 GAGGATTAATGAATGCTGTAGAGACTGTTTTGAAGCTCTCAATACCAATGCTACCTGT 189
Qy 1074 GGCCTGTCATGTTCTATTGCTTTTCCACCTTTTGGTTAAATATCCTGGCAGAGCTTCTTC 1133
Dy 190 GGCCTTGCATGTTTATTGCTTTTCCATCTGTGTTAAACATACTGCTGAGATCTTC 249
Qy 1134 GATTTGGTGCATCGTAATCTACAAAGGATGCTGGAATGCCAAACTGTCGAAGATTATT 1193
Dy 250 GATTTGGTGACCGAGAATCTTACAAAGACTGGTGAATGCAAGACAATTGATGAGTACT 309
Qy 1194 GGAGGATGTGGAATATGCTGTTTCCAAATGGATGATCCGCCACCTATATTTCCATGTT 1253
Dy 310 GGAGAAATGGAACATGCTGTGCATAAATGGATTGTTGTCATATATATTTCCCTTGCA 369
Qy 1254 TAAGGCACGGTCTACCAAGGCTGCTGCTTTTAAATGCTTCTGCTGTTTCTGCTTTAT 1313
Dy 370 TGCGAAATGGTATATCAAGGAAGTGTCTGTTTTTATATCGTCTTCTTCTGCTGTAC 429
Qy 1314 TCCATGAGCTGTGCATTGCT 1333
Dy 430 TTCATGAGCTGCAGATTACT 449

Search completed: February 21, 2003, 16:23:06
Job time : 448 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 16:13:34 ; Search time 2523 Seconds
(without alignments)
12465.967 Million cell updates/sec

Title: US-09-856-018B-15
Perfect score: 1942
Sequence: I tagaaacacgctcgtctt.....aaaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	816.6	42.0	855	10 AW349274	AW349274 GM210004B
c 2	488	25.1	961	14 BQ510367	BQ510367 EST617782
c 3	470.2	24.2	572	13 BM309699	BM309699 sak65f03
4	461.4	23.8	586	14 BQ148998	BQ148998 NF086009F
5	433	22.3	433	9 A1441040	A1441040 sa58f02.y
6	432.2	22.3	729	13 B1422326	B1422326 EST532992

7	427.6	22.0	1181	11	AY110660	AY110660 zea mays
8	424.6	21.9	572	14	BQ124305	BQ124305 EST609881
9	404.8	20.8	1572	11	AY105372	AY105372 zea mays
10	368.2	19.0	741	14	BQ998922	BQ998922 QG620J23
11	334.6	17.2	459	14	BQ123670	BQ123670 EST609246
12	321.8	16.6	539	13	B1422212	B1422212 EST532878
13	307.4	15.8	682	10	BE247899	BE247899 NF038D11D
14	305	15.7	763	13	B1308446	B1308446 EST529856
15	295.8	15.2	729	13	BJ322102	BJ322102 BJ322102
16	295	15.2	577	10	AW586836	AW586836 EST318459
17	290.6	15.0	634	12	BF634363	BF634363 NF059006D
18	288	14.8	540	9	AJ470192	AJ470192 AJ470192
19	282.8	14.6	606	13	BJ316561	BJ316561
20	278.8	14.3	626	10	AV926912	AV926912 AV926912
21	276.8	14.3	561	10	AW035727	AW035727 EST281881
22	255.6	13.2	655	14	BQ998311	BQ998311 QG19C15
23	254.4	13.1	862	12	BG321213	BG321213 Zm04_05g0
24	245.4	12.6	480	9	AJ470191	AJ470191 AJ470191
25	235.6	12.3	480	9	AJ470190	AJ470190 AJ470190
26	238	12.3	528	9	AL381190	AL381190 MCB57D12
27	237	12.2	508	14	BU009402	BU009402 QG10F20
28	233.8	12.0	673	14	BQ998554	BQ998554 QG19N20
29	229.8	11.8	629	9	AA042298	AA042298 24635 CD4
30	220.6	11.4	515	10	AV925760	AV925760 AV925760
31	217.4	11.2	555	13	BJ268713	BJ268713 BJ268713
32	216.8	11.2	685	14	BQ861203	BQ861203 QG17N02
33	212.4	10.9	656	14	BQ862755	BQ862755 QG21P16
34	200.4	10.3	422	12	BF199515	BF199515 WHE0591-0
35	198.6	10.2	617	10	AV930897	AV930897 AV930897
36	180	9.3	531	10	BE356202	BE356202 DGI_123_B
37	178.8	9.2	238	12	BF066952	BF066952 st35e10.y
38	178.6	9.2	706	13	BM348355	BM348355 MEST289-G
39	176	9.1	655	10	AW775077	AW775077 EST334228
40	174.4	9.0	574	12	BF003479	BF003479 EST431977
41	170.8	8.8	701	13	BM334936	BM334936 MEST130-B
42	168.2	8.7	700	13	BM341343	BM341343 MEST333-F
43	164.2	8.5	682	13	BM075161	BM075161 MEST351-B
44	162.6	8.4	681	13	BM267747	BM267747 MEST371-E
45	161.6	8.3	627	10	AV956927	AV956927 AV956927

ALIGNMENTS

RESULT 1
AW349274/c
LOCUS AW349274 855 bp mRNA linear EST 04-OCT-2000
DEFINITION GM210004B21H12 Gm-r1021 Glycine max CDNA clone Gm-r1021-1536 3', mRNA sequence.
ACCESSION AW349274
VERSION AW349274.1 GI:6846984
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 855)
Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Corvelli,V., Erpelting,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
A Functional Genomics Program for Soybean (NSF 9872565)
TITLE Unpublished (1999)
JOURNAL Other ESTs: A1441040
COMMENT Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888)919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genome systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.

[illegible]

```

1. 855
/organism="Glycine max"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="Gm-r1021-1536"
/clone.lib="Gm-r1021"
/tissue_type="root"
/lab_host="XL10-Gold"
/notes=Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
XhoI; Library Gm-r1021 is a sequence-driven, reraoked set
of the original library Gm-cl004 which was prepared from
root cDNA. The mRNA was isolated from entire roots of 8
day old 'Williams' seedlings which were propagated on
paper towels with distilled water. Stratagene's cDNA
Synthesis Kit (catalog #200401) was used to synthesize
the cDNA. The Gm-cl004 library was constructed by Dr.
Paul Keim & Virginia H. Corvelli, Department of Biology,
Box5640, Northern Arizona University, Flagstaff, AZ 86011
, email: paul.keim@uau.edu, virginia.corvelli@uau.edu. The
contig analysis to select unique genes was performed by
the laboratory of Ernest Retzel, Computational Biology
centers, University of Minnesota,
http://www.cbc.umn.edu/Research/Projects/Soybean/Index.html
. Reracking was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and sequencing by the Keck
Center for Comparative and Functional Genomics,
University of Illinois,
http://www.life.uiuc.edu/biotech/keck.html."
170 c 153 q 214 t 31 others

```

```
Query Match          42.0%; Score 816.6; DB 10; Length 855;
Best Local Similarity 95.9%; Pred. NO. 6e-120;
```

Matches	819;	Conservative	0;	Mismatches	35;	Indels	0;	Gaps	0;
QY	1010	AAGGGAACCTCTCTTACGCCACGAGAGAGTCTCTGAAGCTTTCTGTCTCCAAATTTATAT	1069						
Db	854	AAGGAACCTTTNNTTACGCCACNNNNNAGTNNNGAANCNNNNNGTTCCAAATTTATAT	795						
QY	1070	GTGTGGCTCTGCATGTCTATTCGTTTTTCCACCTTTTGGTTAAATATCCTGGCAGAGCTT	1129						
Db	794	CTGTGCTCTGCATGTCTATTCGTTTTTCCNNNNNNNGNTAAATATCCTGGCAGAGCTT	735						
QY	1130	CTTCGATTTGGTGATCGTGAATTTCTACAAGGATTTGGTGAATGCCAAACTGTGCGAAGAT	1189						
Db	734	CTTCGATTTGGTGATCGTGAATTTCTACAAGGATTTGGTGAATGCCAAACTGTGCGAAGAT	675						
QY	1190	TATTTGGAGGATGTGGAATATGCCTGTTCACAAATGGATGATCGCCACCTATATTTTCCA	1249						
Db	674	TATTTGGAGGATGTGGAATATGCCTGTTCACAAATGGATGATCGCCACCTATATTTTCCA	615						
QY	1250	TGTTTAAAGGCACGGCTTACCAAAGGCTGCCTCTTTTAAATTTGGCTTCTCGGTTCTCGCT	1309						
Db	614	TGTTTAAAGGCACGGCTTACCAAAGGCTGCCTCTTTTAAATTTGGNNNNCTGGTTCTCGCT	555						
QY	1310	TTATTTCCATGAGCTGTGCATTTGCTGTTCCTTGGCCACATATTCAGTTGTGGGCTTTCGGT	1369						
Db	554	TTATTTCCATGAGCTGTGCATTTGCTGTTCCTTGGCCACATATTCAGTTGTGGGCTTTCGGT	495						
QY	1370	GGAAATATGTTTCAGGTTCCTTTGGCTTTGATCACTAAATATCTGCAAAATAAATTCAGA	1429						
Db	494	GGAAATATGTTTCAGGTTCCTTTGGCTTTGATCACTAAATATCTGCAAAATAAATTCAGA	435						
QY	1430	AACTCAATGGTTCGGAATAATGATTTTTTGGTTCATATTCAGTATCTCTGGTTCACCTATG	1489						
Db	434	AACTCAATGGTTCGGAATAATGATTTTTTGGTTCATATTCAGTATCTCTGGTTCACCTATG	375						

QY	1490	TGTGTA	CTGCTAT	ACTACCA	TGACTT	GATGA	TAGGA	AGGCA	AACTT	GACTGA	AGCTAC	1549	
Db	374	TGTGTA <th>CTGCTAT</th> <th>ACTACCA</th> <th>TGACTT</th> <th>GATGA</th> <th>TAGGA</th> <th>AGGCA</th> <th>AACTT</th> <th>GACTGA</th> <th>AGCTAC</th> <th>315</th>	CTGCTAT	ACTACCA	TGACTT	GATGA	TAGGA	AGGCA	AACTT	GACTGA	AGCTAC	315	
QY	1550	GGCATT <td>ACACAT</td> <td>TTTAA</td> <td>AGGTG</td> <td>CACAT</td> <td>GATG</td> <td>AGCTT</td> <td>TTTCAG</td> <td>TTTTT</td> <td>CAGATTT</td> <td>GTTAA</td> <td>1609</td>	ACACAT	TTTAA	AGGTG	CACAT	GATG	AGCTT	TTTCAG	TTTTT	CAGATTT	GTTAA	1609
Db	314	GGCATT <td>ACACAT</td> <td>TTTAA</td> <td>AGGTG</td> <td>CACAT</td> <td>GATG</td> <td>AGCTT</td> <td>TTTCAG</td> <td>TTTTT</td> <td>CAGATTT</td> <td>GTTAA</td> <td>255</td>	ACACAT	TTTAA	AGGTG	CACAT	GATG	AGCTT	TTTCAG	TTTTT	CAGATTT	GTTAA	255
QY	1610	ATGTGG	ATATG	TGGTCA	ATAT	TTTGT	TTTCT	ACGAAT	GC	TTTTC	ATCAT	TACCAT	1669
Db	254	ATGTGG	ATATG	TGGTCA	ATAT	TTTGT	TTTCT	ACGAAT	GC	TTTTC	ATCAT	TACCAT	195
QY	1670	CTGCTC	TGAAG	GAAT	CCAC	GGA	TATG	CCAG	GATG	CCAG	GCTT	CAAG	1729
Db	194	CTGCTC	TGAAG	GAAT	CCAC	GGA	TATG	CCAG	GATG	CCAG	GCTT	CAAG	135
QY	1730	TATGTA	CTATT	ACCA	CTCT	CTCT	CGCA	ATTG	TATCA	AAAA	TATG	CAAT	1789
Db	134	TATGTA	CTATT	ACCA	CTCT	CTCT	CGCA	ATTG	TATCA	AAAA	TATG	CAAT	75
QY	1790	CAC	TGGCAT	TGATA	ACTGCC	CAAG	GAAC	ACTCT	AACTG	TTTTT	CTGTT	AACTG	1849
Db	74	CAC	TGGCAT	TGATA	ACTGCC	CAAG	GAAC	ACTCT	AACTG	TTTTT	CTGTT	AACTG	15
QY	1850	AGAGG	CGCTAG	ATGT	1863								
Db	14	AGAGG	CGCTAG	ATGT	1								

RESULT 2

BQ510367/c

LOCUS

EST617782

DEFINITION

Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ76 3' end. mRNA sequence

961 bp

linear

EST 22-JUL-2002

RESULT	2
BQ510367/c	
LOCUS	
DEFINITION	BQ510367 961 bp mRNA linear EST 22-JUL-2002 EST617782 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ76 3' end.. mRNA sequence

3' end, mRNA sequence.
ACCESSION BQ510367
VERSION BQ510367.2 GI:21926064
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliids; eudicotyledons; core eudicots; Asteridae; eustersoids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 961)
AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.
TITLE Generation of a set of potato cDNA clones for microarray analyses
JOURNAL Unpublished (2002)
COMMENT On Jun 10, 2002 this sequence version replaced qi:21369236.

Other ESTs: EST617781
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@tigr.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cdna@resgen.com
Seq primer: T7.

```

FEATURES
source
Location/Qualifiers
1..961
/organism="Solanum tuberosum"
/cultivar="Kennebec or Binjte"
/db_xref="taxon:4113"
/clone="STMJ76"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and phytophthora
infestans-treated libraries of stolons, leaves, leaflets."

```


axillary buds of stem explants, petioles, germinating eyes

BASE COUNT 324 a 205 c 181 g 251 t
ORIGIN

Query Match 25.1%; Score 488; DB 14; Length 961;
Best Local Similarity 75.2%; Pred. No. 5.3e-68;
Matches 627; Conservative 0; Mismatches 195; Indels 12; Gaps 1;

QY 716 TCTATGACATACAACTATGATGAGAGCCTTACCAATTTAGTTGAAAAGGAGAA 775
|||||
DB 961 TCTATGACATACAAATATGATGAGAGCCTTACCAATTTAGTTGAAAAGGAGAA 902
QY 776 GCAGTCTCGATCTGACATGAGTCTTACCAAGTAAAGCTTCAAGAGCTTGGCA 835
DB 901 -----AATTCGGAATCACTCTTACCAATTTAGTTTCAAGAGTTTGGCT 854
QY 836 TATTTCTGTTGCTGCTTACATTTATGTTACCAAGTAAAGCTTCTCGCACCTTATAT 895
DB 853 TACTTCTGATGTTGCTGCTTACATTTATGCTATGCTTACAGTCTGCTGCTGCTGCTAT 794
QY 896 CGAAGGGTTGTTGTTTCCCAACTTGTCAAGCTGATATATTTACAGAGTTATGGGA 955
DB 793 CGAAGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
QY 956 TTTATATACCAAT 1015
DB 733 TTTATCATGAGCAGTAT 674
QY 1016 AACCTTTTACGCCACCCAGAGTCTGCAAGCTTCTGTTCCCAATTTATATGTTG 1075
DB 673 AACCTTTTACGCCACCCAGAGGTTGTTGCAAGCTTCTGTTCCCAATTTATATGTTG 614
QY 1076 CTCTGATGTTCTATGCTTTTTCACCTTGTGTTAAATATCTGTCGAGAGTCTTCTCGA 1135
DB 613 CTCTGATGTTCTACAGCCTTTTCATCTTGGTAAATATATCTGCGGAATTTCTGCGA 554
QY 1136 TTTGGTATCGTGAATTTCTACAGGATTTGGTGAATGCGCAAACTCTGCAAGATTTG 1195
DB 553 TTTGGGATCGTGAATTTCTACAGGATTTGGTGAATGCGCAAACTCTGCAAGATTTG 494
QY 1196 AGGATGGAATATGCTGTTTACAAATGATGATCGCGCACTATATTTTCCATCTTTA 1255
DB 493 AGATTTGGAATATGCTGTTTACAAATGATGATCGCGCACTATATTTTCCATCTTTA 434
QY 1256 AGGACGGTCTACCAAGGCTGCTGCTTTTAAATGCTTCTGTTTCTGCTTTTATTC 1315
DB 433 AGGAATGGCATACCTAAGGAGTTGCAATGGTATCTCTTTTATATCTGCTGTTTC 374
QY 1316 CATGAGCTGATGCTGTTCTTCCACATATTTCAAGTTTGGGCTTTCGGTGGAAAT 1375
DB 373 CATGAGCTATGATGCTGTTCTTCCGCTATTTCAAGTTTGGGCTTTCGGTGGAAAT 314
QY 1376 ATGTTTCAGGTTCTTGTGTTGATCACTAATTTATCTCAAAATAAATTCAGAACTCA 1435
DB 313 ATGTTTCAGATTCCTTGTGATCACTACAGCACTTCTGCAAAACAAAGTTCAAAACTCG 254
QY 1436 ATGTTTGAATATGATTTTTTGGTTTCAATTCAGTATCTTGGTCAACCTATGTTGTA 1495
DB 253 AATGTGGCAACATGACATCTGTTGTTTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTG 194
QY 1496 CTGCTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1549
DB 193 CTCTGTTATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 140

RESULT 3
BM309699
LOCUS
DEFINITION BM309699 sak65f03.y1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID: DIACYLGLYCEROL ACYLTRANSFERASE ;, mRNA sequence.
ACCESSION BM309699

VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORSTITLE
JOURNAL
COMMENTFEATURES
source

BM309699.1 GI:18041405

EST.

soybean.

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

1 (bases 1 to 572)

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corvett,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 412.

Location/Qualifiers

1..572

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1036-7949"

/clone_lib="Gm-cl036"

/tissue_type="Somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This
cDNA library was constructed from mRNA isolated from
somatic embryos (age ranging from 2 months to 9 months)
cultured on MSD 20. The library was prepared using the
Life Technologies pSuperScript cDNA library construction
kit. Complementary DNA was synthesized from mRNA using a
poly (dT) sequence with a NotI restriction site. SalI
linkers adapters were ligated to the blunt-ended cDNA
fragments followed by NotI digestion. The cDNA fragments
were directionally cloned into the NotI-SalI restriction
site of the pSPORT1 vector. The ligated cDNA fragments
were transformed into E.coli Electromax DH10B host cells.
This library was constructed in the laboratory of Dr. Lila
Vodkin by Anu Khanna at the University of Illinois at
Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"BASE COUNT
ORIGIN

Query Match

Best Local Similarity 24.2%; Score 470.2; DB 13; Length 572;

Matches 532; Conservative 0; Mismatches 24; Indels 18; Gaps 3;

QY 55 TGTAGGCACTGCTCTCAACCACTCTTCCCTGCGCGCGCGCTCCCTCCGCCACCTCCACCGC 114

DB 17 TGTACCACTGCTCTCAACCACTCTTCCCTGCGCGCGCGCTCCACCGC-----CCGC 67

QY 115 CGGCTCTTCAATTCGCTGAGACACCAACCGACACTTCGGGTGATGACTTGCCCAAGGA 174

DB 68 TGGCTCTTCAATTCGCTGAGACACCAACCGACACTTCGGGTGATGACTTGCCCAAGGA 127

QY 175 TTCTGTTCCGAGCACTTCCATCAACGACGACGCGCGCTCAATTCGCCACGACAAA 234

DB 128 TTCCGTTCCGAGCACTTCCATCAACGACGACG-----CGCGCAATTCGCCACGCA-- 179

QY 235 CGAAAAAAGACACTGATTTCTCCCTCTCTCAAAATTCGGCTACCGCTTCCTCCGCCG 294

Db 180 -CAAAACAAAGACACTGATTTTCGGTCTCAAAATTCGCCTACCGTCTCCGTCGCCGC 238
 QY 295 TCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGGTCAGAGTCAGCG 354
 Db 239 TCATCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGGTCAGAGTCAGCG 298
 QY 355 GGGCTCTTCAACCTTTGTATAGTAGTCCCTTGTGCTGTGAATAGCCGACTCATCTGA 414
 Db 299 GGGCTCTTCAACCTCTCTATAGTAGTCCCTTGTGCTGTGAATAGCCGACTCATCTGA 358
 QY 415 GAATTTAATGATAGTGTGGTGATCAAAATCTGGCTTTTGGTTAGTTCAAAAGTCATT 474
 Db 359 GAATTTAATGATAGTGTGGTGATCAAAATCTGGCTTTTGGTTAGTTCAAAAGTCATT 418
 QY 475 GAGACTGGCCCTTTTCATGTGTGCTTCTCTGCTGTATTTCTTTTCGGTCCGCTT 534
 Db 419 GAGACTGGCCCTCTTCATGTGTGCTTCTCTGCTGTATTTCTTTTGGTGTGCTT 478
 QY 535 TATAGTGAGAGTGTGGCACACGGAAGTGATACCCGAAACAGTGTGTGTGACTTCA 594
 Db 479 TATAGTGAGAGTGTGGCACACGGAAGTGATACCCGAAACAGTGTGTGTGACTTCA 538
 QY 595 TATATCATTTACCTCAACTTCGCTTTTCTATCCA 628
 Db 539 TATATCATTTACCTCAAGTTCCTACTTTTCTATCCA 572

RESULT 4

B0148998 LOCUS B0148998 586 bp mRNA linear EST 24-APR-2002
 DEFINITION NF086D09FL1F1077 Developing flower Medicago truncatula cDNA clone
 NF086D09FL 5', mRNA sequence.

ACCESSION B0148998
 VERSION B0148998.1 GI:20286057

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM

Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE

1 (bases 1 to 586)
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

Medicago truncatula flower library

COMMENT

Unpublished (2001)

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Insert Length: 586 Std Error: 0.00

Plate: 086 row: D column: 09

Seq primer: TCACACAGAAACAGCTATGAC.

FEATURES

source

1..586

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF086D09FL"

/tissue_type="Developing flower"

/dev_stage="Developmentally pooled"

very young, developing, fully-opened flowers and flowers

in early transition into pods."

/note="Vector: Lambda Zap; cDNA was prepared from polyA+

enriched, pooled samples of equivalent amounts of total

RNA from very young, developing, fully-opened flowers and

flowers transitioning into pods. The cDNA was

directionally ligated into the Uni-Zap XR vector

(Stratagene) and packaged using the Gigapack III Gold
 packaging extracts. Phagemids containing cDNA inserts were
 in vivo excised from the recombinant Uni-Zap XR vector
 using ExAssist helper phage and the E. coli strain
 XL1-Blue MRF (Stratagene). Excised plasmids were plated
 using SOLR cells."

BASE COUNT 134 a 115 c 132 g 202 t 3 others
 ORIGIN

Query Match 23.8%; Score 461.4; DB 14; Length 586;

Best Local Similarity 86.5%; Pred. No. 1.1e-63;

Matches 507; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 837 ATTTCCTGGTGGCCCTTACCATTTATGTTACCAAGCAAGCTATCTCGCACACCTTATATTC 896

Db 1 ACTTCATGGTGGCTTACATATGCTACCAAGCAAGCTATCTCGCACACCTTATGCGTTC 60

QY 897 GAAAGGGTGGTGGTGGCCCAACTTGTCAAGCTGATTAATTTACAGAGAGTTATGGAT 956

Db 61 GAAAGGGTGGTGGTGGTGGCCCAACTTGTCAAGCTGATTAATTTACAGAGAGTTATGGAT 120

QY 957 TTATAATAGACCAATATATTAATCCCATAGTACAAATTCACAGCATCTCTCAAGGAA 1016

Db 121 TTATAATAGACCAATATATTAATCCCATAGTACAAATTCACAGCATCTCTCAAGGAA 180

QY 1017 ACCTTCTTTACGCCACGAGAGAGTTCCTGAAGCTTTCTGTTCCAAATTTATATGTGTGC 1076

Db 181 ACCTTCTATATGCCATTGAGAGAGTTCCTGAAGCTTTCTGTTCCAAATTTATGTGTGC 240

QY 1077 TCTGCATGTTCTATGCTTTTCCACCTTTGGTAAATATATCTCGCAGAGCTTCTTCGAT 1136

Db 241 TGTGCATGTTCTATGCTTTTCCATCTTTGGTAAATATATCTCGCAGAGCTTCTTCGCT 300

QY 1137 TTGTGTGATCGTGAATCTTACAGAGTGTGTGGAATGCCAAACCTGTCGAAGATTTTGA 1196

Db 301 TTGTGTGATCGTGAATCTTACAGAGTGTGTGGAATGCCAAACCTGTCGAAGATTTTGA 360

QY 1197 GGATGTGGAATATGCTGTTTCCAAATGGATGATCCGCGACCTATATTTCCATGTTTAA 1256

Db 361 GGATGTGGAATATGCTGTTGTCACAAATGGATGATGTTGTCGTCACGTGATTTTCCCTGCA 420

QY 1257 GGCAGGCTTACCAAGAGCTGCTGCTTTTAAATGCGCTTCTCGTGTCTGCTTTATTC 1316

Db 421 GGTGTGATATACCAAGAGGCTGCTGCTTTGACTGCTTCTGCTGCTTCTGCTGCTGCT 480

QY 1317 ATGAGCTGTGCATGCTGCTTCTTCCACATATTCAGTGTGGGCTTTCGTCGTAATTA 1376

Db 481 ATGAGTATGATGCTGCTTCTTCCCAATGTCACAGTGTGGGCTTTTATTTGGAATTA 540

QY 1377 TGTTTCAGGTTCTTTTGGTCTTGTATCATTAATTTATCTGCAAAATAA 1422

Db 541 TGTTTCAGGTTCTTTTGGTCTTGTATCATTAATTTATCTGCAAAATAA 586

RESULT 5

AI441040

LOCUS

DEFINITION

AI441040

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max

Eukaryota;

Viridiplantae;

Streptophyta;

Embryophyta;

Tracheophyta;

Spermatophyta;

Magnoliophyta;

eudicotyledons;

core eudicots;

Rosidae;

eurosid 1;

Fabales;

Fabaceae;

Papilionoideae;

Phaseoleae;

Glycine.

REFERENCE

1 (bases 1 to 433)

Shoenmaker, R., Keim, P.,

Vodkin, L., Erpelidg, J.,

Corvelli, V., Khanna

A., Bolla, B., Marra, M.,

Hillier, L., Kucaba, T.,

Martin, J., Beck, C.,

Wylie, T., Underwood, K.,

Thiesing, B., Allen, M.,

Bowers

Y., Person, B., Swaller, T.,

Gibbons, M., Pape, D.,

Harvey, N., Schurk

R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Willson, R.

JOURNAL

Public Soybean EST Project
Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 424

POLYA-No.

FEATURES

Location/Qualifiers

1. 433

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-3508"

/clone_lib="Gm-cl004"

/tissue_type="root"

/lab_host="XL10-Gold"

/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
XhoI; Root cDNA. The mRNA was isolated from entire roots
of 8 day old 'Williams' seedlings which were propagated on
paper towels with distilled water. Stratagene's cDNA
Synthesis Kit (catalog #200401) was used to synthesize the
cDNA. First- strand synthesis was performed with 5-methyl
dCTP, hence the ligated cDNA is hemimethylated.
Stratagene's first-strand synthesis primer was used
[GAGAGAGAGAGAGAGAGAGTAGTCGAG(T)-18]. After
second-strand synthesis, the cDNA ends were 'polished'
with clone pfu DNA polymerase, ligated to EcoRI adapters,
and phosphorylated. The XhoI site within the first-strand
synthesis primer was restricted by digestion with XhoI;
all XhoI sites in the cDNA would be protected by their
hemimethylated status. The cDNA constructs were
size-fractionated with a 500bp cutoff, using GibcoBRL Life
Technologies' cDNA Size Fractionation column. The column
eluent was then ligated into Stratagene's pBluescript II
XR Predigested vector (pBluescript II SK(+)) that had been
digested with EcoRI and XhoI, and phosphorylated. Both
the white and blue colonies appear to contain recombinant
plasmids with cDNA inserts. Blue colonies 9n-15) have been
sequenced, and possess putative cDNA inserts. This library
was constructed by Dr. Paul Keim & Virginia H. Coryell,
Department of Biology, Box 5640, Northern Arizona
University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:
520-523-7500, email: paul.keim@na.u.edu,
virginia.coryell@na.u.edu"

BASE COUNT

102 a 136 c 88 g 107 t

Query Match

Best Local Similarity 22.3%; Score 433; DB 9; Length 433;

Matches 433; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 GAAAGTGTAGCCACGCTCTCAACACACTTTCCTCGCGCGCGCGCTCCCGCCACCTCC 109

DB 1 GAAAGTGTAGCCACGCTCTCAACACACTTTCCTCGCGCGCGCGCTCCCGCCACCTCC 60

QY 110 ACCGGCGGCGCTCTCAATTCGCTGAGACAACACGAGCTTCGCTGATGACTTGGCC 169

DB 61 ACCGGCGGCGCTCTCAATTCGCTGAGACAACACGAGCTTCGCTGATGACTTGGCC 120

QY 170 AAGGATTCGTGGTCCGACGACTCCATCAACAGCGAGCGCGCGCTCAATTCCTCCCAACAG 229

DB 121 AAGGATTCGTGGTCCGACGACTCCATCAACAGCGAGCGCGCGCTCAATTCCTCCCAACAG 180

QY 230 CAAAACGAAAAACAACAGACTGATTTCTCCGCTCAAAATTCGCTACCGCTCCGTC 289
DB 181 CAAAACGAAAAACAACAGACTGATTTCTCCGCTCAAAATTCGCTACCGCTCCGTC 240
QY 290 CCGGCTCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTCCGTCAGAGT 349
DB 241 CCGGCTCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTCCGTCAGAGT 300
QY 350 CACGGCGGCGCTCTCAACCTTTGTATAGTAGTCCTTGTGTGTGAATAGCGGACTCATC 409
DB 301 CACGGCGGCGCTCTCAACCTTTGTATAGTAGTCCTTGTGTGTGAATAGCGGACTCATC 360
QY 410 ATTGAGATTTAATCAAGTAGTGGTTGGTTCATCAATTCGCTTGGTTAGTTCAAG 469
DB 361 ATTGAGATTTAATCAAGTAGTGGTTGGTTCATCAATTCGCTTGGTTAGTTCAAG 420
QY 470 TCATTGAGAGACT 482
DB 421 TCATTGAGAGACT 433

RESULT 6

BI422326

LOCUS

DEFINITION

EST532992 tomato callus, TAMU Lycopersicon esculentum cDNA clone

BI422326

ACCESSION

BI422326

VERSION

BI422326.1

KEYWORDS

EST.

SOURCE

tomato.

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta;

Asteridae; euasterids I; Solanales;

Lycopersicon.

REFERENCE

AUTHORS

Alcala, J., Vrebalov, J., White, R.,

Matern, A.L., Vision, T., Holt, I.E.,

Liang, F., Upton, J., Craven, M.B.,

Bowman, C.L., Ann, S., Ronning,

C.M., Fraser, C.M., Martin, G.B.,

Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES

Location/Qualifiers

1. 729

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="cLE69N10"

/clone_lib="tomato callus, TAMU"

/tissue_type="callus"

/dev_stage="25-40 days old"

/lab_host="xLI-Blue MRF"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons

of seedlings 7-10 days post-germination were excised, cut

at both ends and placed on MS medium with no selection.

Mixed callus was harvested at 25 and 40 days and included

undifferentiated masses. Tomato Callus EST Library"

BASE COUNT

177 a 142 c 160 g 250 t

ORIGIN

Query Match

Best Local Similarity 22.3%; Score 432.2; DB 13; Length 729;

Matches 533; Conservativity 76.0%; Pred. No. 4e-59;

Mismatches 168; Indels 0; Gaps 0;

QY 785 GATACTCGAATCGACTATCTTACACGTAAGCTTCAAGAGCTTGGCATATTTCCG 844

DB 14 GAGAATCCGAATCAACTACTCTTACATGTAGTTTCGAGAGTTGGCTTACTTCATG 73

```
Qy 845 GTTGCCCTACATATGTACAGCAGCTATCCTCGCACACCTTATATTCGAAAGGT 904
Db 74 GTTGCTCAACATTTATGCTATAGCTTATAGCTTCTGCTCGCATTCATTCGAAAGGT 133
Qy 905 TGGTGTTCGCCAACTTGTCAAGCTGATAATATTTACAGGAGTTATGGGATTTATATA 964
Db 134 TGGTGGCCGCCAACTCATCAGCTGTAATTTTAGAGAGTTAATGGGATTTATCAT 193
Qy 965 GACCAATATTAATCCCATAGTACAAATTCACAGCATCTCTCAAGGGAAACCTTCTT 1024
Db 194 GAGCAGTATATTAACCCGATTGTGGAAGCTCACGACATCCATTTGAAGGAAACCTTTTA 253
Qy 1025 TACGCCACGAGAGATTCAGAGCTTCTGTTTCCAAATTTATATGTGTGCTGTGCATG 1084
Db 254 TACGCCATCGAGAGGATTAAGAGCTTTCAGTTCCAAATTTTATATGTGTGCTGTGCATG 313
Qy 1085 TTCTATTGCTTTTCCACCTTTGTTAAATATCTGGCAGAGCTTCTTCGATTTGGTGAT 1144
Db 314 TTCTACAGCTCTTTCATCTTTGGTTAAATATCTTGGGGAAGTTCTGCGAATTTGGGGAT 373
Qy 1145 CGTGAATTTCTACAGGATTTGTTGGAATGCCAAACTGTCTGAAGATTTATGAGGATGTGG 1204
Db 374 CGTGAGTTCTATAAGATTTGTTGGAACGCAAAACAATTTGATGATTTGGAGACTTTGG 433
Qy 1205 AATATGCTGTTCACAAATGATGATCGGCACCTATATTTTCCATGTTTAAAGCAGCTT 1264
Db 434 AATATGCTGTACATAAGTGTGATGTTGCTGCATCATATTTTCCATGCTTAAAGAAATGGC 493
Qy 1265 CTACCAAGGCTGTGCTTTTAAATGCTTCTGCTGCTGCTTATTCATTCAGTACGCTG 1324
Db 494 ATACTTAGGAGTTGCAATGGTGATCTCTTCTTATATCTGCTGTTTCCATGAGCTG 553
Qy 1325 TGCATTTGCTTCTTGCACATATTCAGGTTGCGGCTTTTCGGTGGAAATTTATGTTTTCAG 1384
Db 554 TGTATGCTGTTCTGTCGCTATTCAGGTTTGGGCAATTCCTTGGAAATCATGTTTCAG 613
Qy 1385 GTTCCTTTGCTGTGATCACTAATATCTGGAATTAATTCAGAACTCAATGGTTGGA 1444
Db 614 ATTCCTTTGCTGTACATAAGCAACTTCTGCAAAACAAGTTTCAAAACTCGAATGTGGC 673
Qy 1445 AATATGATTTTGTGTCATATTCAGTATCCTTGTGTCGAACC 1485
Db 674 AACATGACATCTGCTGCTTTTCTGCAATGTTGTCGAACC 714

RESULT 7
LOCUS AV110660 1181 bp mRNA linear HTC 26-MAY-2002
DEFINITION Zea mays CL739_1 mRNA sequence.
ACCESSION AV110660
VERSION AV110660.1 GI:21215250
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1181)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1181)
Coe,E.C.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
LOCATION/Qualifiers
1. .1181
/organism="Zea mays"
/db_xref="MaizeDB:632842"
/db_xref="taxon:4577"
FEATURES
source
```

```
BASE COUNT 322 a 240 c 275 g 327 t 17 others
ORIGIN
/clone="CL739_1"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
```

```
Query Match 22.0%; Score 427.6; DB 11; Length 1181;
Best Local Similarity 69.0%; Pred. No. 1.7e-58;
Matches 587; Conservative 0; Mismatches 261; Indels 3; Gaps 1;

Qy 712 GGTCTCTTATGCACATACAACTATGATATCAGAGCAGCTTACCAAAATAGTTGAAAGGG 771
Db 11 GGTCTCTTATGCACATACAAATTTATGATATAAGGGTATTGTCCAAAAGTACTAGAAAGGG 70
Qy 772 AGAAGCAGCTGCTCGATCTCTGAACATGGACTATCCTTACAACTGATGATATTTACAGAGCTT 831
Db 71 TGCTGCATATGGAATTTATGCTCGATCTCGAATATGAAGATATCAACCTTTAAAGTCT 130
Qy 832 GGCATATTTCTGTTGCCCCCTACATTTATGTTACAGCAGAGCTATCCTCGCACACCTTA 891
Db 131 AGTGTACTTCTATGTTGCCCCCAACACTTTGTTACCCAGCAACTTATCCTCAAACTACATG 190
Qy 892 TATTTCGAAGGGTTGGTTGTTTCCCAACTTGTCAAGCTGATATATTTACAGAGCTTAT 951
Db 191 TATTGAAAGGGTTGGTGGACCCAGCAACTATAAAGTGGTGGNNNNNNACAGGCTTGT 250
Qy 952 GGGATTTTAAATGACCAATATATTAATCCCATAGTACAAATTCACAGCATCTCTCTCAA 1011
Db 251 GGGCTTCATATTCAGCAATATATAAACCCTTGTGAAGAAATTCACCAACATCCACTGAA 310
Qy 1012 GGGAAACCTCTTTACGCCACCGAGAGAGTTCTGAAAGCTTTCTGTTCCAAATTTATATGT 1071
Db 311 AGGGAANNNNNGAATGCTATAGAAAGAGTCTTAAAACTCTCAGTGCACAACTATATATGT 370
Qy 1072 GTGCTCTGCATGTTCTATTCGCTTTTCCACCTTTGGTTAAATATCCTGGCAGAGCTTCT 1131
Db 371 ATGGCTTTGCTATGTTCTATTGCTNNNNNNCATTTTATGCTGAACATTTAGCTGAACCTCT 430
Qy 1132 TCGATTTGCTGATCGTCAATTTCTACAAAGGATTTGGTGAATGCCAAACTGTCGAAGATTA 1191
Db 431 CTGTTTCCGTTGACCGTCAATTCATAGGACTGTTGGAATGCCAAACTGTTGAAGAGTA 490
Qy 1192 TTGGAGGATGGAATATGCTGTTTCACAAATGATGATCGGCCACCTATATTTTCCATG 1251
Db 491 CTGGAGGATGTGAACATGCTGTTTCAATAGTGGATCATCAGACACATATATTTTCCATG 550
Qy 1252 TTTAAGGCACGGTCTACCAAGGCTGCTCTTTTAAATTCCTTCTGCTTCTGCTTCTGCTTT 1311
Db 551 TATAAGGAAAGGCTTTTCCAGGGGTGATGTTTCTTCTTCTGTTTCTGCTTCTGCTGCTG 610
Qy 1312 ATTCCATGAGCTGTGCATTTGCTGTTCCCTGTCACATATTTCAAGTTGTGGGCTTTTCGGTGG 1371
Db 611 ATTCCATGAGATATGATTTGCGGTGCGCTGCCACATTTTCAAAATTCGGGCAATTTCTGG 670
Qy 1372 AATTATCTTTTCAGGTTCTCTTTGGTCTTTGATCACTAATATATCTGCAAAATAAATTCAGAAA 1431
Db 671 GATCATGTTTCAGATACCGTTGGTATTCTTGACAAGATATCTCCATGCTACGTACGTTCAAGCA 730
Qy 1432 CTCATGTTTGGAAATATGATTTTTCGTTTCATATTCAGTATCCTTGGTCAACCTATGTG 1491
Db 731 TGTAAATGGTGGGCAACATGATATTTTGGTCTC---TTCATATAGTTCGGACAGCGGATGTG 787
Qy 1492 TGTACTCTATCTACTTACCATGCTTGTGAATAGGAAGGCAAACTTGAACCTACGG 1551
Db 788 TGTCTCTCTATCTACTTACCATGCTGATGAACAGCGCAGGCCAGGCAAGTAGATAGTTCGG 847
Qy 1552 CCATTACATTT 1562
```

```
Db 848 CAGAGACATGT 858
|||||
RESULT 8
B0124305
LOCUS
DEFINITION EST17-APR-2002
sequence.
ACCESSION B0124305
VERSION B0124305.1 GI:20176267
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 572)
Grusak M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T., Cheung
, F. and Fraser, C.M.
ESTs from late stage developing seeds of Medicago truncatula
Unpublished (2002)
Contact: Grusak, M.A.
USDA/ARS Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713 798 7044
Fax: 713 798 7078
Email: mgrusak@bcm.tmc.edu
TIGR sequence name: MTRBB78TK More information is available at:
www.medicago.org
Seq primer: SKmod (CTA gaa CTA gtg gAT CC).
Location/Qualifiers
1..572
/organism="Medicago truncatula"
/cultivar="A17"
/db_xref="taxon:3880"
/clone_lib="GLSD"
/tissue_type="Immature seeds"
/dev_stage="25 to 35 days after pollination"
/lab_host="XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Immature seeds, collected from pods ranging in age
from 25 to 35 days after pollination, were harvested from
greenhouse-grown plants. Seed were removed and separated
from pod walls and immediately frozen in liquid nitrogen.
Seeds throughout the age range were pooled for mRNA
extraction. cDNA was prepared from polyA+ enriched RNA.
The cDNA was directionally ligated into the Unizap XR
vector from Stratagene and packaged using Gigapack III
Gold packaging extracts. Plasmids containing cDNA inserts
were excised from the recombinant lambda-zap phage using
Ex-assist helper phage and propagated in XLOLR cells."
BASE COUNT 160 a 112 c 105 g 195 t
ORIGIN
Query Match 21.9%; Score 424.6; DB 14; Length 572;
Best Local Similarity 84.9%; Pred. No. 7.1e-58;
Matches 488; Conservative 0; Mismatches 84; Indels 3; Gaps 1;

QY 519 TTCCTTTCGCGCTTTATAGTGAGAAAGTTGGCAACAGGAGTGTATACCGAACACG 578
|||||
Db 1 TTCCTTTCGCGCGCTTTATATACAAAGTTGGCAACAGGAAAGTGTATATGAAACATA 60
|||||
QY 579 TTCTTTTGTACTTCATATAAATCAATACCTCACTTCGCTTTTCTATCCAGTTTACTTA 638
|||||
Db 61 TTGTTTGTCTATATATATCAATTCATTCGCAACCACTCGAGTCTCTATCTCTATTTAGTAA 120
|||||
QY 639 TTCTCAGGTGTATCTGCTTTGTATCAGGTGTACAGTTTATGCTGTTTCTCTGTGTTG 698
|||||
Db 121 TCCTCAGGTGTATCTGCTTTGTGTCAGGCGTCACATTCATGCTACTGCTTGCATTG 180
|||||
```

```
Qy 452 TTTTGGTTTGTAGTTCAAGTCAATTTGAGAGACGGCCCTTTTCATGTTGTGCTTTCTCTT 511
Db 1 TTTTGGTTTGTAGTTCAATTTGAGAGACGGCCCTTTTCATGTTGTGCTTTCTCTT 60
Qy 512 GTGGTATTTCTTTCGCTGCTTTATAGTGGAGAGTTGGCACAAGGAAAGTGTATACCC 571
Db 61 CCCATATTTCCCTTGGTGGCAATTTGCACTGAAAGTTGGCATTTCAACAACTCTCGTTAGT 120
Qy 572 GAACCAAGTTGTTGTACTTCAATATATACCTCAACTCGCTTTTCTATCCAGTT 631
Db 121 GATCTGCTACTACCTGTTTTCACATCTTTTACAACTTTGAAATTTGATATCCAGTG 180
Qy 632 TTAGTTATTTCTCAGTGTGATTTCTGCTTTTGTATCAGTGTGACGTTTAACTGCTGTTTCT 691
Db 181 CTCGTGATTTCTTAAGTGTGATTTCTGCACTTTTATCAGGCTTTTGTGATGTTTATTCGC 240
Qy 692 TGTGTTGTATGTTTAAATTTGGTGTCTTTATGCAATACATAAATGATATGAGAGCACTT 751
Db 241 TGCATTTGTTGGCTGAAGCTTTGTATCTTTTGCACATACAAACCATGATATAGAAGAACTG 300
Qy 752 ACCAAATTTAGTTGAAAGGGAGAGCACTGCTGATACCTCTGAACATGCACTATCCCTTAC 811
Db 301 ATCACAAGCGCAGAGAGTTGATATGAACCTGACCGCGCTGGCATAGATAATTTACAA 360
Qy 812 AACGTAAGCTTCAAGAGCTTGGCATATTTCTGTTGGTCCCTACATATGTTTACCGCCA 871
Db 361 GCTCCAACCTTTGGGAGTCAACATATCTCATGATGCTCGACACTCTGTTATCAGCCA 420
Qy 872 AGCTATCTCCACACCTTATATTCGAAGAGTTGGTGTGTTTCCGCCAACTTGTCAAGCTG 931
Db 421 AGTTATCTCCCAACCTTATGTTAGAAAGGTTGGTGTGCTGCTCAAGTTTATCTATAC 480
Qy 932 ATAATATTACAGAGTTATGGGATTTATATAGACCAATATATTAATCCCATAGTACAA 991
Db 481 TTGATATTTACTGCTCCAGAGTTCTATATGAGCAATACATAATCCCTATTTGTTG 540
Qy 992 AATTCACAGCATCTCTCAAGGAAACCTTCTTACGCCACCGAGAGTTCTCAAGCTT 1051
Db 541 AACTCTCAACATCCATGATGGAGGATTACTGAATGCTGTAGAGACTGTTTTGAAGCTC 600
Qy 1052 TCTGTTCCTCAATTTATATGCTGCTGCTGATGCTTCTATTTCTTTCCACTTTGGTTA 1111
Db 601 TCATTTACCAATGCTACCTGTGGCTTTGCAAGTTTATTTGCTTTTCCATCTGTGGTTA 660
Qy 1112 AATATCTGCGAGAGCTTCTTCGATTTGGTGTGATGCTGAATTTCTACAGGATTTGGGAAT 1171
Db 661 AACATCTTGTGAGATTTCTTCGATTTGGTGGACCGAGAAATTTCAAGAGACTGGTGGAT 720
Qy 1172 GCCAAACTGTCGAAGATTTATGGAGGATGGAATATGCTGTTTCAAAATGATGATC 1231
Db 721 GCAAGACAATTTGATGAGTACTGGAGAGAAATGGAACATGCTGTGCATATAATGATGTT 780
Qy 1232 CGCCACCTATATTTTCCATGTTTAAAGCACGCTTACCAAGGCTGCTGCTCTTTTAAAT 1291
Db 781 CGTATATATATTTTCTTTCATGCAAGAAATGATATCAAAAGGAAGTGTGCTTTTATA 840
Qy 1292 GCCTTCTCTGTTTCTGCTTTATTCATGAGCT 1323
Db 841 TCGTTCTTGTGCTGCTACTTCATGAGCT 872

RESULT 10
BQ998922
LOCUS
DEFINITION BQ998922 741 bp mRNA linear EST 22-AUG-2002
            OGS20123.yg ab1 OG-EFGHJ lettuce serriola Lactuca sativa cDNA clone
            OGS20123, mRNA sequence.
ACCESSION BQ998922
VERSION BQ998922.1 GI:22433318
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
AUTHORS
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,
P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_CA_Contig2177, see http://cgdb.ucdavis.edu/
for details.
Plate: OGS20 row: J column: 23.
FEATURES
.....
source
1..741
/organism="Lactuca sativa"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="OGS20123"
/clone_lib="OG-EFGHJ lettuce serriola"
/lab_host="E.coli"
/notes="Vector: pBRCDNA51AB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgdb.ucdavis.edu/
TAG LIB=OG-EFGHJ lettuce serriola
TAG TISSUE=flowers pre-fertilized
TAG_SEQ=GCTTGCAGGG"
BASE COUNT 197 a 145 c 144 g 255 t
ORIGIN
Query Match 19.0%; Score 368.2; DB 14; Length 741;
Best Local Similarity 70.3%; Pred. No. 5.3e-49;
Matches 507; Conservative 0; Mismatches 213; Indels 1; Gaps 1;
Qy 267 AATTTCGCTACCGTCTTCCGTCACGAGTACGAGAGTCTTTCACCTTTGTATAGTCTCTG 326
Db 20 ATTATGCGTATCGACCGTCTTCTCCAGTCTATCGAGAAATTAAGAGTCTCTCTAAGTT 79
Qy 327 CCGACACTATTTTCGTCAGAGTACGCGGGCCCTTTCACCTTTGTATAGTCTCTG 386
Db 80 CTGACGCCATTTTCAAGCAGAGTCTATGAGACTCTTTAACTTTTGCATAGTCTCTAG 139
Qy 387 TTGCTGTGAATAGCGGACTCATCTATGAGAAATTAATGAAGTATGTTGGTTGATCAAT 446
Db 140 TTGCACTCAATGAGGACTCATCTATGAGAAATTAATGAAGTATGTTGGTTGATCAAT 199
Qy 447 CTGCTTTTGGTTTGTGTTAGTCAAGTATGAGAGACTGGCCCTTTTTCATGTTGTCTTT 506
Db 200 CCAATTTTGGTTTCAAGTATGAGAGTATGAGAGTATGAGAGTATGAGAGTATGAGAGT 259
Qy 507 CTCTTTGTTGTTTTCCTTTCGCTGCTTATAGTGGAGAGTGGCACAAGGAGTGTATA 566
Db 260 CTCITCCAAATCTTCCCGTTACTGCTCTATATTTGTTGAAATTAAGTATGAGAGTAT 319
Qy 567 TACCCGAACCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 626
Db 320 TTTTCAGACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 379
Qy 627 CAGTTTGTAGTTTCTCAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 686
```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.

1 (bases 1 to 741)

Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison, P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compenomics.ucdavis.edu/

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]

belongs to contig OG_CA_Contig2177, see http://cgdb.ucdavis.edu/ for details.

Plate: OGS20 row: J column: 23.

Location/Qualifiers

1..741

/organism="Lactuca sativa"

/cultivar="L.serriola"

/db_xref="taxon:4236"

/clone="OGS20123"

/clone_lib="OG-EFGHJ lettuce serriola"

/lab_host="E.coli"

/notes="Vector: pBRCDNA51AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/ TAG LIB=OG-EFGHJ lettuce serriola TAG TISSUE=flowers pre-fertilized TAG_SEQ=GCTTGCAGGG"

/organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLEC69M10"
 /clone_lib="tomato callus, TAMU"
 /tissue_type="callus"
 /dev_stage="25-40 days old"
 /lab_host="X1-Blue MRF"
 /note="vector: pBluescript SK(-); Site1: EcoRI; Site2:
 XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
 of seedlings 7-10 days post-germination were excised, cut
 at both ends and placed on MS medium with no selection.
 Mixed callus was harvested at 25 and 40 days and included
 undifferentiated masses. Tomato Callus EST Library"
 BASE COUNT 137 a 102 c 120 g 180 t
 ORIGIN

Query Match 16.68; Score 321.8; DB 13; Length 539;

Best Local Similarity 75.8; Pred. No. 1.3e-41; Mismatches 127; Indels 0; Gaps 0;

785 GATACCTCGAATGAGTATCTTACAACTGAAGCTTCAAGAGCTTGGCATATTTCTCG 844

14 GAGATTCGGAATCACTACTCTTACAACTGTAGTTTCGAGAGTTGGCTTACTCATG 73

845 GTTGCCCTACATATGTTACCAAGCAAGCTATCCTCGCACACCTTATATTCGAAAGGGT 904

74 GTTGCTCAACTTTATGCTATCAGTATAGCTATCTCTGCTGTCATCCATTCGGAAGGGT 133

905 TGGTGTTCCTCCCACTGTGCAAGCTGATATATTTACAGGAGTTATGGGATTTATATA 964

134 TGGCTGGCCGCCCACTCATCAAGCTGGTAATTTTACAGGATTAATGGGATTTATCAT 193

965 GACCAATATATTAATCCATAGTACAAATTCACAGACTCTCTCAAGGGAACCTTCTT 1024

194 GAGCAGTATATTAACCGATTGTGGAAGCTCAGACATCCATTGAAGGAACCTTTTA 253

1025 TACGCCACCGAGAGATTCCTGTTTCCAAATTTATATGTTGCTCTGCATG 1084

254 TACGCCATCGAGAGGATTTGAAGCTTTCAGTTCCAAATTTTATATGTTGCTCTGCATG 313

1085 TTCTATTGCTTTTCCACCTTTGTTAAATATCTGCGAGAGCTTCTCGATTGTTGAT 1144

314 TTCTACAGCTCTTTTCATCTTTGTTAAATATATCTGCGGAAGTTTTCGGAATTTGGGGAT 373

1145 CGTGAATTTCAAGAGATTTGTTGAATGCCAAACTGTGGAAGATTTATTTGAGGATGTGG 1204

374 CGTGATTTCTAAGATTTGTGGAACCGCAAAACAATTGATGATGATTTGAGACTTTGG 433

1205 AATATGCTGTTCACAAATGATGATCGGCCACCTATATTTTCCATGTTTAAAGCAGGTT 1264

434 AATATGCTGTACATAAGTGGTGGTTCGTCACATCTATTTCCCATGCTTAAAGAAATGGC 493

1265 CTACCAAGGCTGCTGCTCTTTTAAATGCTTCCTGCTGCTTCTGCT 1309

494 ATACCTAAGGAGTGTGAATGGTATCTCTTTTATATCTGCT 538

RESULT 13

BE247899

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

BE247899

VERSION

BE247899.1

KEYWORDS

EST.

SOURCE

ORGANISM

barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

REFERENCE 1 (bases 1 to 682)

AUTHORS

Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula drought library

Unpublished (2000)

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Medicago Genome Initiative accession: MGI:S:25730

Insert Length: 682 Std Error: 0.00

Plate: 038 row: D column: 11

Seq primer: TCACACAGAAACAGCTATGAC.

Location/Qualifiers

1. 682

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone_lib="Drought"

/tissue_type="Plantlets"

/dev_stage="Pooled timepoints"

/note="Vector: Lambda Zap; Contains a mixture of entire

plantlets harvested in a series of days-post-watering

timepoints."

BASE COUNT 169 a 158 c 148 g 207 t

ORIGIN

Query Match 15.88; Score 307.4; DB 10; Length 682;

Best Local Similarity 73.3; Pred. No. 2.3e-39;

Matches 467; Conservative 0; Mismatches 151; Indels 19; Gaps 5;

QY 13 TCGGTCTCTCTCCATGCGGATTTCCGATGAGCCTTGAAGAGTAG-----CCAC 63

Db 50 TTGTTTTTATTTCTCCGATGGCGATTTCCGAGACCTTCACGGCCGCTCCGGAACCCAAA 109

QY 64 TGCCTCTCAACACATCTTC---CCTGGCGCGCGCTCCCTCCGCCACCTCCACCGCGCGCT 120

Db 110 GGTCTCAACGACTCTCTGTTTACGTCCGAGCGCGCGCGCTTCGTCACGGGCTT 169

QY 121 CTTCAATTCGCTGAGAC---AACCACGACAGTTCGGGTGATGACTTGGCCAAAGGATTC 177

Db 170 CTTGGAATTCGCTGCCACCGGTGGAAAGCGAGATTCGGGTGAAGAGTTCGGTGAAGGATTC 229

QY 178 TGGTTCGAGACTCCATCACGAGCGCGCGCTCAATTCCTCAACAGCAACAGCA 237

Db 230 TAGTTCCGATGACTCTATCAGCAGTACCAAAACATCGCTGCTGCC---AATTCGATCA 286

QY 238 AAAACAAGACACTGATTTCTCCGCTCCTCAATTCGCTACCGTCTTCCTCCGTCGCCGCTCA 297

Db 287 AATCGAGGGCTGATATTTCTCGATCAATTCACATTACCGGCTTCTGTTCGCCGCTCA 346

QY 298 CCGCAAGAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGTCAGATCAGCGGG 357

Db 347 TCGTAGAATCAAGGAAAGTCCACTCAGCTCGATAACATTTTCGTCAGAGTCAATCTGG 406

QY 358 CCTCTCAACCTTTGTATAGTAGTCTGTTGTGCTGAATAGCCGACTCATCTTGAAGAA 417

Db 407 CCTCTCAACCTTTGTATAGTAGTCTGTTGTGCTGAATAGCCGACTCATCTTGAAGAA 466

QY 418 TTTAATGAAGTATGGTGGTGTGATCAATCTGCTTTTGGTTTGTAGTTTCAAGTCAATTTAG 477

Db 467 TTTAATGAAGTATGGTGGTGTGATCAATCTGCTTTTGGTTTGTAGTTTCAAGTCAATTTAG 526

QY 478 AGACTGCGCCCTTTTCATGTTGTTCTTTCTTTTGTGTTGTTTTCCTTTTCGCTGCTTTAT 537

Db 527 AGATTGGCCCTCTTCTATGTTGTTCTTAGTCTTAGTCTTAGTCTTAGTCTTAGTCTTAGT 586

QY 538 AGTGGGAAGTGGCAACAGGAGTCTATACCGAACAGTCTTCTTGTGTACTTTCATAT 597

Db 587 AGTGGAAAGTGGCAACAG- GAAGTGTATTTATGACATATATGTTGTCTACTTTCATATA 645

QY 598 AATCATTAACCTAAGTTCGTTTTCATCCAGTTTAA 634
 Db 646 TAATCATTCACAACTCCAGTTCTCTATCTCTATTTAA 682

RESULT 14 BI308446

LOCUS BI308446 763 bp mRNA linear EST 20-JUL-2001
 DEFINITION EST529856 GPOD Medicago truncatula cDNA clone pGPOD-5N2 5' end,
 mRNA sequence.

ACCESSION BI308446

VERSION BI308446.1 GI:14982773

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE 1 (bases 1 to 763)
 AUTHORS Grusak M.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T., Cho
 J., and Fraser, C.M.

TITLE ESTs from developing reproductive tissues of Medicago truncatula
 JOURNAL Unpublished (2001)

COMMENT Contact: Michael A. Grusak
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA
 Tel: 713-798-7044
 Fax: 713-798-7078
 Email: mgrusak@bcm.tmc.edu
 B395301e

TIGR sequence name: MTOAP73TK
 More information is available at: www.medicicago.org
 Seq primer: SKmod (CTA gaa cta gtg gat CC).

FEATURES

source

1..763

/organism="Medicago truncatula"

/cultivar="Al7"

/db_xref="taxon:3880"

/clone="pGPOD-5N2"

/clone_lib="GPOD"

/tissue_type="Immature pod walls"

/dev_stage="Immature pods, ranging in age from 15 to 30

days after pollination"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; Immature pods, ranging in age from 15 to 30 days

after pollination, were collected from greenhouse-grown

plants. At harvest, seeds were removed from pods and

isolated pod walls were collected and immediately frozen

in liquid nitrogen. Pod walls were pooled for mRNA

extraction. cDNA was prepared from polyA+ enriched RNA.

The cDNA was directionally ligated into the Unizap XR

vector from Stratagene and packaged using Gigapack III

Gold packaging extracts. Plasmids containing cDNA inserts

were excised from the recombinant lambda-zap phage using

Ex-assist helper phage and propagated in XL049 cells."

BASE COUNT 198 a 174 c 155 g 236 t

ORIGIN

Query Match

Best Local Similarity 15.7%; Score 305; DB 13; Length 763;

Matches 444; Conservative 0; Mismatches 135; Indels 18; Gaps 4;

QY 13 TCGGTCTTCTCCAAATGGGATTTCCGATGCGCTGAAGTGTAG-----CCAC 63

Db 170 TTGCTTTTATTTCCGATGGGATTTCCGACGACCTCAGGGCGCTCGGAACCCAA 229

QY 64 TGCTCTCAACCACTCTTC---CCTGGCGCGCGTCCCTCCGCGACCTCCACCGCGCGCT 120

Db 230 GGTCTCAACGACTCTCTGTTTACGTGCGAGCGGCGGACGCGCTCCGTAGCGGCTT 289

QY 121 CTTCAATTCCGCTGAGAC---AACCAACCAGACAGTTCGGTGATGACTTGGCCAAAGATTTC 177
 Db 290 CTTGGAATCGTCTGCCACGGTGGAAAGCGAGAGTTCGGTGGAAGATCGGTGAAGGATTC 349
 QY 178 TGGTTCGAGCACTCCATCAACAGCGACGACGCGCGCTCAATTCCCAACAGCAAAACGA 237
 Db 350 TAGTTCGAGTACTCTATCAGCAGTGACCAAAACATCGCTGCTGCC--AAATCTGATCA 406
 QY 238 AAAACAAGACACTGATTTCTCCGTCCTCAAAATTCGCTACCTCGTTCGCTCCGCTCCCTCA 297
 Db 407 AAATCGAGGCGCTGATATTTCTTCGATCAAAATTCACCTACCGCGCTCTGTTCCCGCTCA 466
 QY 298 CGCAAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTCCCGTCAGAGTACCGGG 357
 Db 467 TCGTAGAATCAAGAAAGTCCACTCAGCTCCGATAACATTTCCCGTCAGAGTCATCTGG 526
 QY 358 CTTCTTCAACCTTTGTATAGTAGTCTTGTGTAATAGCGGACTCATCATTCAGAA 417
 Db 527 CTTCTTCAACCTCTGTATAGTAGTCTTGTGTAATAGCGGACTCATCATTCAGAA 586
 QY 418 TTTAATGAAGTATGTTGTTGATCAAAATCTGGCTTTTGGTTTAGTTCAAAGTCATTTAG 477
 Db 587 TTTAATGAAGTACGGTTGTTGATTAAGACTGGCTTTTGGTTTAGTTCAAATTCATTTAG 646
 QY 478 AGACTGGCGGCTTTTCATGTTGTTCTCTCTGTTGGGATTTTCCTTCGTCGCTTTAT 537
 Db 647 AGATTGGCGGCTTTTCATGTTGTTCTCTCTGTTGGGATTTTCCTTCGTCGCTTTAT 706
 QY 538 AGTGGAGAAGTTGGCAACAGGAAGTGTATACCGCAACAGTTGTTGTTGTTACTTCA 594
 Db 707 AGTGGAAAAGTGGCAACAGGAAGTGTATTTATGAACATATTTGGTGGTCTACTTCA 763

RESULT 15

BI322102/c

LOCUS BI322102

DEFINITION

aestivum cDNA clone whyf23117 3', mRNA sequence.

ACCESSION BI322102

VERSION BI322102.1

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 729)

AUTHORS Ogiwara, Y. and Murai, K.

TITLE Expressed genes in Triticum aestivum

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..729

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="whyf23117"

/tissue_type="spikelet at early flowering"

/dev_stage="Feekes' scale 6"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site_1: EcoRI; Site_2: XhoI; Plants were grown under

hydroponic conditions at UC Davis, salt stressed for 12

hours, and for 7 days, then dissected and frozen (Akhunov

in J. Dvorak Lab). Total RNA was prepared from sheath

tissue, equal quantities of RNA were pooled from the two

samples, polyA was purified from the pooled RNA, a cDNA

library was made, and the cDNA clones were in vivo excised to give pbluescript phagemids in the TJ Clouse lab at the University of California, Riverside (Akhunov, Chin, Choi, Clouse, Pention, Kianian, Otto, Simons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT
ORIGIN

237 a 146 c 153 g 193 t

Query Match 15.2% Score 295.8; DB 13; Length 729;
Best Local Similarity 68.7%; Pred. No. 1.5e-37;
Matches 436; Conservative 0; Mismatches 197; Indels 2; Gaps 2;

Qy 914 CGCCAACCTTGCAAGCTGATAATTTACAGGAGTTATGGGATTTATAATAG-ACCAATA 972
Db 729 CGTCAAAATTTCTGCACTTGCATATTACAGGTATTCAAAGGTTTCATTATTGAAGCAGTA 670

Qy 973 TATTAAATCCCATAGTACAAAATTC-ACAGCATCTCTCAAGGGAACCTCTTTACGCCA 1031
Db 669 CATAAATCCCAATCGTCGTAACCTCCCAACATCTCAAGGCGGACTTCTGAATGCTG 610

Qy 1032 CCGAGAGAGTTCTGAAGCTTCTCTCCAAATTTATATGCTGCTGTCATGTTCTATT 1091
Db 609 TAGAGACTGTTCTGAGGCTCTCATACCAATGTTTACTTATGGCTTTGCAATGTTCTATT 550

Qy 1092 GCTTTTCCACCTTTGGTTAAATATCCTGGCAGAGCTTCTTCGATTTGGTGATCGTGAAT 1151
Db 549 GCTTTTCCATCTCTGGTTAAATATCTGCTGAGATTTCTCGTTTCGGTGACCGTGAAT 490

Qy 1152 TCTACAAGGATGCTGGNAATGCCAAACTGCGAAGATTATTGGAGGATGCGGAATATGC 1211
Db 489 TCTACAAGACTGGTGGAAATCAAAACAAATGATGAGTATTGGCGAAATGGAACATGC 430

Qy 1212 CTGTTCAAAATGGATGATCGCCACCTATATTTTCCATGTTTAAGGCACGGCTACCAA 1271
Db 429 CCGTGCATAAATGGATGTTGCTCATATATATTTTCCCTTGCATCGGAAGTGGTATATCGA 370

Qy 1272 AGGCTGCTGCTCTTTAAATGCGCTTCTGCTGTTTATTCATCAGCTGTGCAATG 1331
Db 369 AGGAAGTTGCTGTTTGTATCATTTTGTATCTGCGCTGCTCCATGAGCTTGTGTTG 310

Qy 1332 CTGTTCTCTGCCACATATTCAGTTGCGGCTTTTCGGTGGAAATTAATGTTTCAGGTTCTTT 1391
Db 309 CTGTCCTCTGCCGAATTCCTGAAGTCTGGGCAATTCCTAGGGATCATGTCGAGATCCCC 250

Qy 1392 TGGTCTTGATCAGTAAATATCTGCAAAATAAATTCAGAACTCAATGTTGGAATATGA 1451
Db 249 TTATCAGATTGACATCATGCTCAAAAGCAAAATTCAGGGATACAAATGCCGCGCAACATGA 190

Qy 1452 TTTTGTGGTTCATATTCAGTATCCTTGGTCAACCTATGTGTGTACTGTACTATPACTACCATG 1511
Db 189 TATTTTGGTCTTTTCTGCTATCTACGCCAGCCCTATGTGGGTTCTCTCTACTACCATG 130

Qy 1512 ACTTGATGAATAGCAAGGCAAACTTGACTCAAGC 1546
Db 129 ATGTGATGAACAGGATTTGGGAAGACAGAATAGAGC 95

Search completed: February 21, 2003, 18:31:26
Job time : 2564 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 16:12:13 ; Search time 95 Seconds
(without alignments)
6269.119 Million cell updates/sec

Title: US-09-856-018B-15
Perfect score: 1942
Sequence: 1 tagaacaacgcctcgtctt.....aaaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA:*
- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	719.8	37.1	1942	US-09-326-203A-1	Sequence 1, Appli
2	229.8	11.8	629	US-09-103-754A-3	Sequence 3, Appli
3	199.4	10.3	234	US-09-326-203A-3	Sequence 3, Appli
4	194.6	10.0	275	US-09-326-203A-5	Sequence 5, Appli
5	187.2	9.6	267	US-09-326-203A-4	Sequence 4, Appli
6	170.2	8.8	1650	US-09-103-754A-2	Sequence 2, Appli
7	162.6	8.4	1976	US-09-165-042-2	Sequence 2, Appli
8	159.8	8.2	1766	US-09-326-203A-15	Sequence 15, Appli
9	159.8	8.2	1766	US-09-326-203A-16	Sequence 16, Appli
10	159.4	8.2	1895	US-09-326-203A-14	Sequence 14, Appli
11	124.4	6.4	253	US-09-326-203A-7	Sequence 7, Appli
12	114.2	5.9	254	US-09-326-203A-8	Sequence 8, Appli
13	108.2	5.6	262	US-09-326-203A-9	Sequence 9, Appli
14	92	4.7	325	US-09-326-203A-10	Sequence 10, Appli
15	65.6	3.4	4011	US-08-121-057-3	Sequence 3, Appli
16	65.6	3.4	4011	US-08-509-187D-3	Sequence 3, Appli
17	65.6	3.4	4011	US-09-121-396-3	Sequence 3, Appli
18	65.6	3.4	4011	PCT-US93-0970A-3	Sequence 3, Appli
19	65.6	3.4	4079	US-08-121-057-2	Sequence 2, Appli
20	65.6	3.4	4079	US-08-509-187D-2	Sequence 2, Appli
21	65.6	3.4	4079	US-09-121-396-2	Sequence 2, Appli
22	65.6	3.4	4079	PCT-US93-0970A-2	Sequence 2, Appli
23	60.4	3.1	3238	US-08-123-934A-5	Sequence 5, Appli
24	60.4	3.1	3238	PCT-US94-10080-5	Sequence 5, Appli
25	60	3.1	1582	US-08-545-196B-10	Sequence 10, Appli
26	60	3.1	1582	US-08-545-196B-12	Sequence 12, Appli
27	59.2	3.0	2025	US-09-149-476-316	Sequence 316, Appli

28 59.2 3.0 2296 4 US-08-496-841C-137 Sequence 137, App

29 59.2 3.0 2346 4 US-09-149-476-193 Sequence 193, App

30 59 3.0 2040 3 US-09-165-042-4 Sequence 4, Appli

31 58.6 3.0 943 4 US-09-149-476-35 Sequence 35, Appli

32 58.2 3.0 746 4 US-09-013-810-1 Sequence 1, Appli

33 58 3.0 289 1 US-08-341-568-3 Sequence 3, Appli

34 58 3.0 289 2 US-08-911-020-3 Sequence 3, Appli

35 57.2 2.9 1225 1 US-08-286-020-1 Sequence 1, Appli

36 57.2 2.9 1225 1 US-08-603-919-1 Sequence 1, Appli

37 56.8 2.9 3437 3 US-08-860-339-17 Sequence 17, Appli

38 56 2.9 1313 4 US-09-149-476-112 Sequence 112, App

39 55.8 2.9 1117 4 US-09-247-373B-33 Sequence 33, Appli

40 55.4 2.9 2469 4 US-09-111-730-5 Sequence 5, Appli

41 55.2 2.8 1143 3 US-09-149-476-119 Sequence 119, App

42 55.2 2.8 2550 6 5258287-23 Patent No. 5258287

43 54.6 2.8 519 4 US-09-227-357-76 Sequence 76, Appli

44 54.4 2.8 1534 1 US-08-300-903A-6 Sequence 6, Appli

45 54.4 2.8 3214 1 US-08-484-105-17 Sequence 17, Appli

ALIGNMENTS

RESULT 1

US-09-326-203A-1

; Sequence 1, Application US/09326203A

; Patent No. 6444876

; GENERAL INFORMATION:

; APPLICANT: Lasser, Mike

; APPLICANT: Ruzinsky, Diane

; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic

; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO

; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 1942

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-326-203A-1

Query Match 37.1%; Score 719.8; DB 4; Length 1942;

Best Local Similarity 73.2%; Pred. No. 7.6e-163;

Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;

Qy 241 ACAAGACACACTGATTTCTCCGTCCTCAAAATTCGCCTACCGTCTCCGTCCTCCCGTCACCG 300

Db 521 AGAAGGAGAGAGAAAGCCGATGCTACGTTTACGTATCGACCGTCGGTTCAGCTCATCG 580

Qy 301 CAAAGTGAAGAAAGTCGGTCAGCTCGACACTATTTTCGGTCAGAGTACAGCGGGCCT 360

Db 581 GAGGCGGAGAGAGAGTCCACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 640

Qy 361 CTTCACCTTTGTATAGTAGTCCCTTTGTGTGTAAGCCGACTCATCATTTGAGATTT 420

Db 641 ATTCACTCTGTGTAGTAGTCTTATTCGTGTAAACAGTAGACTCATCATCGAAATCT 700

Qy 421 AATGAAGTATGTTGGTGTATCAAAATCTGGCTTTTGGTTTGTAGTTCAAAAGTCATTGAGAGA 480

Db 701 TATGAGTATGTTGGTGTATGATCAGACCGATTTCTGGTTTGTAGTTCAAGATCGCTGCGAGA 760

Qy 481 CTGGCCCTTTTCANGTGTCTTCTTCFTGTGGTATTTCTTCGCTGCTTTATAGT 540

Db 761 TTGGCGCTTTTCATGTTGTGTATATCCCTTCGATCTTCTTCTTTGGCTGCTTTACGGT 820

Qy 541 GGAGAGTGGCAACCGAAGTGTATACCGACCGAGCTTGTGTGTGTACTTCATATAT 600

```
Db 821 TGAGAAATGGTACTTCAGAAATACATATACAGAACCTGTGTGCATCTTCTTCATATAT 880
QY 601 CATTACCTCAACCTTCGCTTTCTATCCAGTTTATAGTATATCTCAGGTGTGATCTCGTTT 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 TATCACCATGACAGAGTTTGTATCCAGTTTACGTACCCCTAAGGTGTGATCTCGTTT 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 TGTATCAGGTGTACAGCTTAAATGCTGTTTCTGTGTGTGTATGTTTAAATTTGGTGTCTTA 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 TTTATCAGGTGTACCTTGTATGCTCCTCACTTGCATTTGTGCTAAAGTTGTTTCTTA 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 721 TGCACATACAAATATGATAGACACTTACCAATAGTTTGAAGGGGAGAACACT 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 TGTCTATAGTAGTATGACATAAGATCCCTAGCCAAATGACGTGATGAAG----- 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 781 GCTCGATACTCTGAACATGACTATCCTTACACAGTAAAGCTTCAAGAGCTTGGCATATTT 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1051 -----CCAATCCTGAAGTCTCCTACTAGTTAGCTTGAAGAGCTTGGCATATTT 1099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 841 CTTGTTGGCCCTTACATATATATGTTTACAGCAAGCTATCTCGCACACCTTATATTCGAA 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 CATGGTGGCTCCACATTTGTATCAGCAAGTTATCCAGCTTCTGCATGTATACGGAA 1159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 901 GGTGTGTTGTTTCGCAACTTGTCAAGCTGATAATATTTACAGGAGTTATGGATTTAT 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 GGTGTGTTGTTTCGCAACTTGTCAAGCTGATAATATTTACAGGAGTTATGGATTTAT 1219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 961 ATAGACCAATATATTAATCCCATAGTACAAATTCACAGCATCTCTCAAGGGAACCT 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1220 ATAGAACATATATATATATCTTATCTCAGAACTCAAGCATCTTGAAGGCGATCT 1279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1021 TCTTTACGCCACGAGAGAGTCTGAAGCTTCTGTGTTTCCAAATTTATATGTGTGCTCTG 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1280 TCTATATGCTATTAAGAGAGTGTGAAGCTTTCAGTTCCTCAAAATTTATATGTGTGCTCTG 1339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1081 CATGTTCTATGCTTTTCCACTTGTGTAATATATCTTCCGAGAGCTTCTCGATTGG 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1340 CATGTTCTATGCTTTTCCACTTGTGTAATATATCTTCCGAGAGCTTCTCGATTGG 1399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1141 TGATCGTGAATCTCAAGAGTGTGGAATGCCAAACTGTGGAAGATTTATGGAGGAT 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1400 GGATCGTGAATCTCAAGAGTGTGGAATGCCAAACTGTGGAAGATTTATGGAGGAT 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1201 GTGGAAATGCTGTTACAAATGGAATGATCGGCCACCTATATTTTCCATGTTTAAAGCA 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1460 GTGGAAATGCTGTTACAAATGGAATGATCGGCCACCTATATTTTCCATGTTTAAAGCA 1519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1261 CGGTCTACCAAGGCTGCTCTTTTAAATGCTTCTGCTTCTGCTTTATTCATGA 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1520 CAAGATACCAAGACACTCGCCATTAATATGCTTTTCTAGTCTCTGCACTCTTTTATGA 1579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1321 GCTGTGATTTGCTTCTTCCATATATTCAGATTTGTGGGCTTTTCGGTGGATTTATGTT 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1580 GCTATGATCGGAGTCTTCTGCTCTCTTCAAGCTATGGCTTTTCTTGGGATTTATGTT 1639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1381 TCAGGTTCTTGGCTTGTATCACTAATTTATCTGCAAAATAAATTCAGAAACTCAATAGT 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1640 TCAGGTTCTTGGCTTGTATCACTAATTTATCTGCAAAATAAATTCAGAAACTCAATAGT 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1441 TGAATATGATTTTGTGTTTATATTCATATTCATCTTGTGCTGCTTGTGCTGCT 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1697 GGGAAATGATTTTGTGTTTATATTCATCTTGTGCTGCTTGTGCTGCTGCTGCT 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1501 ATACTACCATGACTTGTGAATAGGAAGGCAAACTTTGACTGAAGCTACGG 1551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1757 TTATTACCAAGCTGATGAACGGAAGGATCGATGTCATGAACAACTG 1807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 2

US-09-103-754A-3

; Sequence 3, Application US/09103754A

; Patent No. 6344548

; GENERAL INFORMATION:

; APPLICANT: Farese, Robert

RESULT 3

US-09-326-203A-3

```
APPLICANT: Cases, Sylvaine
APPLICANT: Smith, Steven
APPLICANT: Erickson, Sandra
TITLE OF INVENTION: Diacylglycerol O-acyltran
TITLE OF INVENTION: sferase
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Bozicevic & Reed
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/103,754A
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 6510-105p
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650 327 3400
TELEFAX: 650 327 3231
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-103-754A-3
```

Query Match 11.8%; Score 229.8; DB 4; Length 629;
Best Local Similarity 79.6%; Pred. No. 6.2e-46;
Matches 284; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

QY 888 CTTATATTCGAAGGGTGTGGTTCGCGCAACTTGTCAAGCTGATATATTACAGGAG 947
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 3 CATGTATACGGAAGGGTGTGGTTCGCGCAACTTGTCAAGCTGATATATTACAGGAG 62
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 948 TTATGGGATTTTATATAGACCAATATATTATCCCATAGTACAAAATTCACAGCATCTC 1007
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 63 TCATGGGATTTTATATAGACCAATATATTATCCCATAGTACAAAATTCACAGCATCTC 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1008 TCAAGGAAACCTTCTTTACGCCACCGAGAGTTCCTGAAGCTTTCTGTTCCTCAAAATTTAT 1067
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 123 TGAAGGGCATCTCTCATATGCTATTGAAAGAGTGTGGAAGCTTTTCAGTTCCAAATTTAT 182
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1068 ATGTGTGGCTTCGCATGTTCTATTGCTTTTCCACCTTTTGGTTAAATATCTCGGAGAGC 1127
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 183 ATGTGTGGCTTCGCATGTTCTATTGCTTTTCCACCTTTTGGTTAAATATCTCGGAGAGC 242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1128 TTCTTCGATTTGTTGATCGTGAATCTACAGGATTTGGTGGAACTGCGAACTGTGGAAG 1187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 243 TTCTCTGCTTCGGGATCGTGAATCTACAAAGATTTGGTGAATGCAAAAGTGTGGGAG 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1188 ATTA-TTGAGGATGTGGAATATGCTTTCACAAATGGATGATCGGCCACCTATAT 1243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 303 ATTACTGGGAGATGTGGAATATGCTTTCCTCCATATGATGATGGTGGTCCGACATATAT 359
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

; Sequence 3, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:

; APPLICANT: Lassar, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 234

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (24)..(60)

; OTHER INFORMATION: n at positions 24, 41, and 60 is unknown

US-09-326-203A-3

Query Match 10.3%; Score 199.4; DB 4; Length 234;

Best Local Similarity 95.7%; Pred. No. 7.8e-39;

Matches 225; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 815 GTAAGCTTCAAGAGCTTGCCATATTTCTGGTTGCCCTA-CATTATGTACCAGCCAAAG 873

DB 1 GTAAGCTTCAAGAGCTTAGCATANTCTCTGGTTGCCCTTANCAATTATGTACCAGCCAAAN 60

QY 874 CTATCTCTGCACACCTTATATTCGAAAGGTTGGTTTTCGCCAACTTGTCAAGCTGAT 933

DB 61 CTATCTCTGCACACCTTATATTCGAAAGGTTGGTTTTCGCCAACTTGTCAA-CTGAT 119

QY 934 ATATTTACAGGAGTATGGGATTTATATAGACCAATATATTAATCCCATAGTACAAA 993

DB 120 AATATTTACAGGAGTATGGGATTTATATAGACCAATATTAATCCCATGTACAAA 179

QY 994 TTCACAGCATCTCTCAAGGGAACTTCTTTACGCCACCGAGAGAGTCTGAAG 1048

DB 180 TTCACAGCATCTCTCAAGGGAACTTCTTTACGCCACCGAGAGAGTCTGAAG 234

RESULT 4

US-09-326-203A-5

; Sequence 5, Application US/09326203A

; Patent No. 6444876

; GENERAL INFORMATION:

; APPLICANT: Lassar, Mike

; APPLICANT: Ruzinsky, Diane

; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic

; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO

; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 275

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (192)..(263)

; OTHER INFORMATION: n at positions 192, 202, 204, 211, 222, 234, 238,

; OTHER INFORMATION: 239, 244, 245, 247, 251, 262, and 263 is unknown
US-09-326-203A-5

Query Match 10.0%; Score 194.6; DB 4; Length 275;

Best Local Similarity 82.2%; Pred. No. 1.2e-37;

Matches 226; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 1165 GTGGAATGCCAAACCTGTCGAAGATTTATGGAGATGTGGAATATGCTTGCCTTCAAAATG 1224

DB 1 GTGGAATGCCAAACCTGTTGAAGATTTATGGAGATGTGGAATATGCTTGCCTTCAAAATG 60

QY 1225 GATGATCCGCCACCTATATTTCCATGTTTAAGGACCGCTTACCACCAAGGCTGCTCTCT 1284

DB 61 GATGATCCGCCACCTATATTTCCATGTTTAAGGACCGCTTACCACCAAGGCTGCTCTCT 120

QY 1285 TTTAATTTGCTTCTCCTGGTTTCTGCTTTATTTCCATGAGCTGTGCATTTGCTTCTTGGCCA 1344

DB 121 TTTAATTTGCTTCTCCTGG-TTCTGCTTTATTTCCATGAGCTGTGCATTTGCTTCTTGGCCC 179

QY 1345 CATATTTCAAGTTTGGGCTTTCGGTGGGAATATGTTTCAAGTTCCTTTGCTTGTATGATCAC 1404

DB 180 ACATATTTCAAGTNGTGGTTTCNGNGGAATTNAGTTTCAGTNCCTTGGGTTTCNACCN 239

QY 1405 TAATTTCTGCAAAATATAATTCAGAACTCAATGG 1439

DB 240 AATTNNNGGCNAAAAAATTCNNGAACCCCGGG 274

RESULT 5

US-09-326-203A-4

; Sequence 4, Application US/09326203A

; Patent No. 6444876

; GENERAL INFORMATION:

; APPLICANT: Lassar, Mike

; APPLICANT: Ruzinsky, Diane

; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic

; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO

; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 267

; TYPE: DNA

; ORGANISM: Glycine max

US-09-326-203A-4

Query Match 9.6%; Score 187.2; DB 4; Length 267;

Best Local Similarity 83.6%; Pred. No. 6.8e-36;

Matches 224; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 654 CTGCTTTTGTATCAGGTGTCAGTTAAATGCTTTTCTTGTGTGTATGTTTAAATTTGG 713

DB 1 CTGCTTTTGTATCAGGTGTCAGTTAAATGCTTTTCTTGTGTGTATGTTTAAATTTGG 60

QY 714 TGTCTTATGCACATACAACTATGATATGAGACGCTTACCATAATAGTTGAAAGGGAG 773

DB 61 TGTCTTATGCACATACAACTATGATATGAGACGCTTACCATAATAGTTGAAAGGGAG 120

QY 774 AAGCACTGCTGATCTCTGAACATGGACTATCCTTACACGCTTACAGCTTCAAGAGCTTGG 833

DB 121 AAACATTACCCCAATCTTTG-ATATGGAGTATCCGTACACTGTGACCTTCAGAGTTGG 179

QY 834 CATATTTCTGTTGCCCTTACATATGTTTACGACCAAGCTATCCTCCGACACCTTATA 893

DB 180 CATACTTCATGTTGCTCTCTACATATGCTATCAGACAGATGATCTCTCGACACCTTCAG 239

QY 894 TTCGAAAGGGTGGTTGTTTCGCCAACT 921

```
|||||
Db 240 TTCGAAGGGTTGGGTGTTCTGCTCAACT 267

RESULT 6
US-09-103-754A-2
; Sequence 2, Application US/09103754A
; Patent No. 6344548
; GENERAL INFORMATION:
; APPLICANT: Farese, Robert
; APPLICANT: Cases, Sylvaine
; APPLICANT: Smith, Steven
; APPLICANT: Erickson, Sandra
; TITLE OF INVENTION: Diacylglycerol O-acyltran
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,754A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 6510-105p
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650 327 3400
; TELEFAX: 650 327 3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1650 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-103-754A-2

Query Match 8.8%; Score 170.2; DB 4; Length 1650;
Best Local Similarity 56.2%; Pred. No. 1.6e-31;
Matches 341; Conservative 0; Mismatches 263; Indels 3; Gaps 1;

QY 798 TGGACTATCCCTTACAACTGTTCAAGAGCTTGGCATATTTCTGTTGCCCTCAT 857
|||
Db 850 TGAGCTATCCAGACAACCTGACCTACCGAGATCTATTACTTCTTCTCTCTT 909
|||
QY 858 TATCTTACCACCAAGCTATCTCGCACACCTTATATTCGAAAGGGTTGGTTTCCGCC 917
|||
Db 910 TGTGTTATGAAGTCAACTTCTCTCGGTCGCCGCAATACGAAAGCGCTTTCTGTTACGAC 969
|||
QY 918 AACCTGTGCAAGCTGATATATTTACAGAGTTATGGGATTTATATACACCAATATATTA 977
|||
Db 970 GAGTCTTGAGTCTCTTTTACCAGCTTCAAGTGGGCTGATCCACAGTGGATGG 1029
|||
QY 978 ATCCCATAGTACAAAATTCACAGCATCCTCTCAAGGAA---ACCTTCTTTACGCCACCG 1034
|||
Db 1030 TCCCTACTATCCACAACCTCCATGAAGCCCTTCAAGGATATGAGCATATTCACGGATCAT 1089
|||
QY 1035 AGAGAGTCTGAAAGCTTTCGTTCCAAATTTATATGTTGCTGCTGCTGCTATTTGCT 1094
|||

Db 1090 AGCGTCTCTTAAAGCTGGCGGTCCCAACCATCTGATCTGGCTTATCTTCTTATGGT 1149
|||
QY 1095 TTTTCCACCTTTGGTTAAATATCTGGCAGAGCTTCTTTCGATTTGGTGATCGTGAATTTCT 1154
|||
Db 1150 TTTTCCACTCTCTGCTCAATGCTGTGGCAGAGCTTCTGCACTTTGGAGACCGGAGTTCT 1209
|||
QY 1155 ACAAGGATTTGGTGAATGCCAAACTCTGCGAAGATTTATGGAGGATGTGGAATATGCTGT 1214
|||
Db 1210 ACAGAGATTTGGTGAATGCTGAGTCTGTCACCTACTTTTGGCAGAACTGGAATATATCCCG 1269
|||
QY 1215 TTCACAAATGATGATCCGCCACCTATATTTTCCATGTTTAAAGGACAGGCTACCAAAAGG 1274
|||
Db 1270 TGCACAAATGATGATCAGACACTTCTACAAAGCCATGCTCAGACATGGCAGCAGCAAT 1329
|||
QY 1275 CTGCTGCTCTTTAAATTCCTTCTGTTTCTGTTTATTCATGAGCTGTCATTTGCTG 1334
|||
Db 1330 GGTGGCCAGCAGAGGAGTATTTTGGACCTCAGCCTTCTCCATGAGTACCTAGTGAGCG 1389
|||
QY 1335 TTCCTTGGCCACATATTCAAAGTTGTGGGCTTTCGGTGGAAATATATTTTCAGGTTTCTTTGG 1394
|||
Db 1390 TTCCCTCGCGATGTTCCGCCCTCTGGGCATTCACAGCCATGATGCTCAGCTCCCACTGG 1449
|||
QY 1395 TCTTGAT 1401
|||
Db 1450 CCTGGAT 1456
|||

RESULT 7
US-09-165-042-2
; Sequence 2, Application US/09165042
; Patent No. 6100077
; GENERAL INFORMATION:
; APPLICANT: Sturley, Stephen L.
; APPLICANT: Oelkers, Peter
; TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL
; FILE REFERENCE: 0575/56331
; CURRENT APPLICATION NUMBER: US/09/165,042
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Yeast
; US-09-165-042-2

Query Match 8.4%; Score 162.6; DB 3; Length 1976;
Best Local Similarity 55.1%; Pred. No. 1.1e-29;
Matches 340; Conservative 0; Mismatches 274; Indels 3; Gaps 1;

QY 789 CTCTGAACATGGACTATCTTACACGTTAAGCTTCAAGAGCTTGGCATATTTCTTCTGGTTG 848
|||
Db 957 CGCACCGTGAGTACCCGGACAATCTGACCTACCGGATCTCTACTACTTCTCTCTTCG 1016
|||
QY 849 CCCCTACATTATGTTACACGCCAAGCTATCTCGCACACCTTATATTCGAAAGGGTTGGT 908
|||
Db 1017 CCCCCACCTTGTGCTACGAGCTCAACTTCCCGCTCTCCCCGATCCGGAAGCGCTTC 1076
|||
QY 909 TGTTTCGCCAACTTGTCAAGCTGATAATATTTACAGAGTTATGGGATTTATATAGACC 968
|||
Db 1077 TGTCTGCGACGGATCTTGGATGCTGTTCTTCAACCCAGCTCCAGGTGGGCTGATCCAGC 1136
|||
QY 969 AATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGAA---ACCTTCTTT 1025
|||
Db 1137 AGTGGATGGTCCCCCATCCATCCAGAACTCCATGAAGCCCTTCAAGGACATGACATCTAC 1186
|||
QY 1026 ACGCCACCGAGAGAGTTCTTGAAGCTTTTCTGTTCCAAATTTATATCTGTGGCTTCTGCA 1085
|||
Db 1197 GCATCATCGAGCGCTCTCTGAAGCTGGGGTCCCAATACCTCATCTGCTCATCTTCT 1256
|||
QY 1086 TCTATTGCTTTTCCACCTTTGGTTAAATATTCCTGGCAGAGCTTCTTCGATTTGGTGATC 1145
|||
```

Db	1100	CCACTCATGTCTCAATGCTGTGGCAGAGCTCTGCAGATTTGGAGACCGGAGTTCTACAG	1159
Qy	1159	GGATTGGTGGAAATGCCAAAACCTGTCGAAGATTATTGGAGGATGTGGAATATGCCTGTTC	1218
Db	1160	GGACTGGTGGAAATGCTGAGCTGTCACTACTTTTGGCAGAACTGGAATATCCCGTGCA	1219
Qy	1219	CAAAATGGATGATGCCGCCACTATATTTTCATGTTTAAAGCAGCGTACCAAAAGCGTC	1278
Db	1220	CAAGTGGTGATCAGACACTTCTACAAGCCTATGCTCAGATGGGCAGCAACAATGGAT	1279
Qy	1279	TGCTCTTTAAATGGCTTCCTGGTTCTGCTTTATTCATCAGCTGTGCATGCTGTTC	1338
Db	1280	GGCCAGGACTGGGGTCTTTTGGCGTCAGCCCTCTTCATCAGTACCTAGTAGGACATTCC	1339
Qy	1339	TTGCCACATATTCAAGTTCTGGGCTTTGGTGGAAATTATGTTTCAGGTTCTCTTTGGTCTT	1398
Db	1340	CTTGAGGATGTTCCGGCTCTGGCAATTACAGCCATGATGCGTCAGGTCGCCACTGGCGTG	1399
Qy	1399	GAT 1401	
Db	1400	GAT 1402	

RESULT 9

```

US-09-326-203A-16
; Sequence 16, Application US/09326203A
; Patent No. 644876
; GENERAL INFORMATION:
; APPLICANT: Lassner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nuclei
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-326-203A-16

```

Query Match	8.2%	Score 159.8	DB 4	Length 1766
Best Local Similarity	55.2%	Pred. No. 4.9e-29		
Matches 333	Conservative 0	Mismatches 267	Indels 3	Gaps 1
QY 802	CTATCCTTACAAAGTAACGCTTCAGAGCCTTGGCATAATTTCTCTGGTTGGCCCTACATATTAG	861		
Db 800	CTATCCGGACAACTGACCTACCGAGATCTCTATTACTTTCATCTTTGTCTCTACTTTGTG	859		
QY 862	TTACAGCGCAAGCTATCTCTCCACACCTTATATTCGAAAGGCTTGTTGTTTCGCGCAACT	921		
Db 860	TTATGAACCTCACTTTCTCTCATCCCCCGGAATCGCAAGGCGCTTTCGTCTACGGCGGGT	919		
QY 922	TGTCAGCGTGATAATATTTACAGGAGTTATGGGATTTATAATAGACCAATATATTAATCC	981		
Db 920	TCTTGAGATGCTCTTTTTCACCCAGCTTCAAGTGGGGCTGATCCAGCAGTGGATGTCCTC	979		
QY 982	CATAGTACAAATTCACAGCATCTCTCAAGGAA ---ACCTTCTTTTACGCCACCGAGAG	1038		
Db 980	TACTATCCAGAACTCCATGAAGCCCTTCAAGGACATGGACTATTCAACGAATCATTTAGCGC	1039		
QY 1039	AGTTCTCAAGCTTCTGTGTTCCAAATTTATATCTGTGGCTCTCGAATGTTCTATTGCTTTT	1098		
Db 1040	TCTCTTAAGCTGGGGTCCCCAACCAATCTGATATGGCTCATCTTCTTCTATTGCTTTT	1099		
QY 1099	CCAGCTTTGGTTAAATATPCTCTGGCAGAGCTCTCTCGATTTGGTGATCGTGAATCTCAAA	1158		

Query Match	8.2%;	Score 159.8;	DB 4;	Length 1766;
Best Local Similarity	55.2%;	Pred. No. 4.9e-29;		
Matches 333;	Conservative 0;	Mismatches 267;	Indels 3;	Gaps
QY	802	CTATCTCTACAAGCTTAAGCTTCAAGAGCTTGCCATATATTTCTGGTGGCCCTACATTTATG	861	
Db	800	CTATCCGGACAACCTGACCTACCGAGATCTCTATTACTTCAATCTTGTCTCTACTTTGTG	859	
QY	862	TTACCAGCCAAAGCTATCTCTCGCACACCTTATATTTCGAAGGGTGGTTGTTTCCGCAACT	921	
Db	860	TTATGACTCAACTTTCCTCGATCCCGCCGAATACGAAGCGCTTTCGCTACGCGGGT	919	
QY	922	TGTCAGCTGATAAATATTTACAGAGGTATGGATTTATAATAGACCAATATATTAATCC	981	
Db	920	TC TTGAGATGCTCTTTTTCACCCAGCTTCAAGTGGGGCTGATCCAGCAGTGGATGGTCCC	979	
QY	982	CATAGTACAAAATTCACAGATCCCTCTCAAGGGAA---ACCTTCTTTACGCCACCGAGAG	1038	
Db	980	TACTATCCAGAATCTCATGAAGCCCTTCAAGGACATGGACTATTCCAGCAATCATTTGAGCG	1039	
QY	1039	AGTTCTCAAGCTTTCGTTCGAAATTTATATCTGTGGCTCTGCCATGTCTAATTCCTTTTT	1098	
Db	1040	TC TCTTAAAGCTGGCGGTCCCCAACCATCTGATATGGCTCATCTTCTTCTAATTGGCTTTT	1099	
QY	1099	CCACCTTTGGTTAAATATCCTGCGAGAGCTTCTTCGATTTGGTGATCGTGAATCTACAA	1158	

Db 1100 CCACTCATGTCTCAATGCTGGGAGAGCTCCTGCAGTTTGGAGACCGCGAGTTCTACAG 1159
QY 1159 GGATTGGTGAATGCCAAAGCTTCGAAGATTATTGGAGGAGTGGGAATATGCCCTGTTC 1218
Db 1160 GGACTGGTGAATGCTGAGTCTGTCACTACTTTGGCAGAGACTGGGAATATGCCCGGTGCA 1219
QY 1219 CAAATGAGTATCGCCGACCTATATTTTCCATGTTTAAAGGCACGCTCTACAAAGGCTGC 1278
Db 1220 CAACTGGTGCATCAGACACTTCTCAAGGCTATGCTCAGACTGGGACGACAAATGGAT 1279
QY 1279 TGCCTTTTAAATGCCCTCCTGGTTCTGCTTTTATTCATGAGCTGTGCATGCTGTTC 1338
Db 1280 GGCCAGGACTGGGCTCTTTTGGGCTCAGCCTTCTCCATGAGTACCTAGTGAACATTC 1339
QY 1339 TTGCCACATATCAAGTTGTGGGCTTTCGGTGGAAATATGTTTCAGGTTCTCTTGGTCTT 1398
Db 1340 CCTGAGGATGTTCCGCTCTGGGCAATTCACAGCCATGATGCTCAGGTCCACCTGGCCTG 1399
QY 1399 GAT 1401
Db 1400 GAT 1402

RESULT 10

US-09-326-203A-14
; Sequence 14, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Lassner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1895
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (209)
; OTHER INFORMATION: n at position 209 is unknown
US-09-326-203A-14

Query Match 8.2%; Score 159.4; DB 4; Length 1895;
Best Local Similarity 54.8%; Pred. No. 6.3e-29;
Matches 338; Conservative 0; Mismatches 276; Indels 3; Gaps 1;
QY 789 CTCGTAACATGGAGTATCCTTACAACCTTAAGCTTCAAGAGCTTGGCATATTTCTGTTG 848
Db 658 CGCACACCGTGAGCTACCCGAGCAATCTGACCTACCCGATCTCTACTTCTCTCTCG 717
QY 849 CCGCTACATTATGTTACGACGAGCTATCTCTGCACACCTTATATTTGAAAGGTTGGT 908
Db 718 CCCCCACCTTGTGTAGAGCTCAACTTCCCGGCTCTCCCGCATCCGGAAGCGCTTC 777
QY 909 TGTTCGCCCACTGTCAAGCTGATATATTTACAGGAGTATGGGATTATATATACACC 968
Db 778 TGCTGACGAGGATCTTGAGATGCTGTCTTCACCCAGCTCCAGGTGGGCTGATCAGC 837
QY 969 AATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAA---ACCTTCTTT 1025
Db 838 AGTGGATGGTCCCAACCATCCAGAACTCCATGAAGCCCTTCAAGGACATGGACTACTCAC 897
QY 1026 ACGCCACCGAGAGTCTCGAAGCTTCTGTTCCAAATTTATATGTTGGCTCTGCAATG 1085

Db 898 GCATCATCGAGCGCTCCTGAAGCTGGCGTCCCAATACCTCATCTGGCTCATCTTCT 957
QY 1086 TCTATTCTTTTCCACCTTTGGTTAAATATCCCTGGCAGAGCTTCTTCGATTTGGTGATC 1145
Db 958 TCTACTGGCTTCTCCATCTCCTGCTGAATGCCGTGGTGGCTCATGCAGTTTGGAGACC 1017
QY 1146 GTGAATCTACAAGGATTTGGGAATGCCAAAGCTGTGCAAGATTTATTGGAGGATGTGA 1205
Db 1018 GGGAGTTCTACCGGAGCTGGTGGAACTCCGAGTCTGTCACTACTTCTGGCAGAACTGGA 1077
QY 1206 ATATGCTGTTTCAAAATGGAATGATGCCGCCACCTATATTTTCCATGTTTAAAGGCACGTC 1265
Db 1078 ACATCCCTGTGCACAAAGTGGTGCATCAGACACTTCTACAAGCCCCATGCTTCGACGGGCA 1137
QY 1266 TACCAAGAGCTGCTGCTCTTTTAAATTCGCTTCCCTGGTTTCTGCTTTTATTCATGAGCTGT 1325
Db 1138 GCAGCAAGTGGATGGCCAGGACAGGGGTGTTCCTGGCTCGGCCCTTCTCCACGAGTACC 1197
QY 1326 GCATTGCTTCTTGGCCACATATTTCAAGTTGTGGGCTTTCGGTGGAAATATGTTTCAGG 1385
Db 1198 TGGTGAAGCTCCCTCTCGGAATGTTCCGCTCTGGGCTTCACGGGCATGATGCTCAGA 1257
QY 1386 TTCCTTTGGCTTGATC 1402
Db 1258 TCCCACTGGCTGGTTC 1274

RESULT 11

US-09-326-203A-7
; Sequence 7, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Lassner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326,203A
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 253
; TYPE: DNA
; ORGANISM: Zea mays
US-09-326-203A-7

Query Match 6.4%; Score 124.4; DB 4; Length 253;
Best Local Similarity 73.2%; Pred. No. 6.9e-21;
Matches 186; Conservative 0; Mismatches 56; Indels 2; Gaps 2;
QY 1196 AGGATGGGAATATGCTGTTTCAAAATGGATGATCCGACCTATATTTTCCATGTTTA 1255
Db 1 AGAAATGGAAATGCTGTCGATAAATGGATGTTGCTCATATATATTTTCTTGGCATG 60
QY 1256 AGGACGCTTACCAAGGCTGCTCTTTTAAATTCGCTTCTGCTTTCTGCTTTATTC 1315
Db 61 CGAAATGGTATATCAAGGAAGTTGCTGTTTATATATCGTT-CTTGTTCCTGCTACTT 119
QY 1316 CATGACCTGTGCATGCTGTTTCTTCCACATATTCAGTTTGGGCTTTCGGT-GGAAT 1374
Db 120 CATGAGTTATGTTGCTGTTTCTTCCCTGCCACATCTCAAGTCTTGGGCTTTTTTAGGAAT 179
QY 1375 TATGTTTTCAGTCTCTTGTGCTTGTGATCAATATCTCAGAAATAAATTTAGAAATC 1434
Db 180 CATGCTTCAGATTTCCCTTCATCATATTTGACATCATACCTCAAAATAAATTCAGTGACAC 239
QY 1435 AATGGTTGGAATA 1448

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 15:48:04 ; Search time 5015 Seconds
(without alignments)
11269.721 Million cell updates/sec

Title: US-09-856-018b-15
Perfect score: 1942
Sequence: 1 tagaaaaacactcgtgttt.....aaaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_nam.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	810.2	41.7	1964	8	AF298815	AF298815 Perilla f
2	754	38.8	2090	8	AY084052	AY084052 Tropaeolu
3	744.6	38.3	2099	8	AF129003	AF129003 Nicotiana
4	719.8	37.1	1904	8	ATH238008	ATH238008 Arabidops
5	719.8	37.1	1942	6	AX090345	AX090345 Arabidops
6	719.8	37.1	1942	6	AX090349	AX090349 Sequence
7	719.8	37.1	1988	8	ATH131831	ATH131831 Arabidops
8	719.8	37.1	2005	8	AY054480	AY054480 Arabidops
9	715	36.8	1845	8	AF051849	AF051849 Arabidops
10	701.2	36.1	1537	8	AF251794	AF251794 Brassica
11	700.2	36.1	1512	8	AF164434	AF164434 Brassica
12	605.6	31.2	1446	8	AF155224	AF155224 Brassica
13	199.4	10.3	234	6	AX090350	AX090350 Sequence
14	194.6	10.0	275	6	AX090352	AX090352 Sequence
15	187.2	9.6	267	6	AX090351	AX090351 Sequence
16	170.2	8.8	1650	10	AF078752	AF078752 Mus muscu
17	170.2	8.8	1776	10	BC003717	BC003717 Mus muscu
18	162.6	8.4	1792	9	AF236018	AF236018 Cercopith
19	162.6	8.4	1976	9	AF059202	AF059202 Homo sapi
20	161.4	8.3	3498	3	CBRG33E23	AC084551 Caenorhab
21	159.8	8.2	728	10	RNO345014	AJ345014 Rattus no
22	159.8	8.2	1751	10	AF296131	AF296131 Rattus no
23	159.8	8.2	1766	6	AX090340	AX090340 Sequence
24	159.4	8.2	1493	9	BC006263	BC006263 Homo sapi
25	159.4	8.2	1895	6	AX090339	AX090339 Sequence
26	159.4	8.2	1998	9	BC015762	BC015762 Homo sapi
27	153	7.9	2074	3	AY051835	AY051835 Drosophil
28	152.2	7.8	1935	4	AY093657	AY093657 Sus scrof
29	151.4	7.8	1497	3	AF221132	AF221132 Caenorhab
30	151.4	7.8	2117	3	AF468650	AF468650 Drosophil
31	151.4	7.8	2342	3	AF468649	AF468649 Drosophil
32	151.4	7.8	11966	3	CEH19N07	292835 Caenorhabd
33	137.6	7.1	253305	3	PFMAL3P7	AL034559 Plasmodiu
34	124.4	6.4	253	6	AX090354	AX090354 Sequence
35	114.2	5.9	254	6	AX090355	AX090355 Sequence
36	108.4	5.6	645	11	PMIC2B	AL684529 Penicilli
37	108.2	5.6	262	6	AX090356	AX090356 Sequence
38	101	5.2	92822	8	AC005917	AC005917 Arabidops
39	101	5.2	97495	-8	AC003058	AC003058 Arabidops
40	92	4.7	325	6	AX090357	AX090357 Sequence
41	88.4	4.6	396	6	AX045795	AX045795 Sequence
42	86.2	4.4	262797	3	CNS07EGG	AL590449 chromosom
43	78.4	4.0	185932	2	AP003714	AP003714 Oryza sat
44	75.2	3.9	3024	10	CGU47320	U47320 Cricetulus
45	71.6	3.7	40494	2	AC005122	AC005122 Drosophil

ALIGNMENTS

RESULT 1	AF298815	Perilla frutescens diacylglycerol acyltransferase (Dgat1) mRNA,	1964 bp	mRNA	linear	PLN 16-OCT-2000
LOCUS	AF298815	complete cds.				
DEFINITION	AF298815					
ACCESSION	AF298815.1	GI:10803052				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
		Perilla frutescens.				
		Perilla frutescens				
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
		Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.				
REFERENCE		1 (bases 1 to 1964)				
AUTHORS		Hwang,S.-K. and Hwang,Y.-S.				

TITLE Isolation of *Perilla frutescens* diacylglycerol acyltransferase cDNA

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1964)

AUTHORS Hwang, S.-K. and Hwang, Y.-S.

JOURNAL TITLE
Direct Submission
 Submitted (23-AUG-2000) Division of Biochemistry, NIAST, RDA, 249
 Seodun-dong, Gwonsun-gu, Suwon 441-707, Korea

FEATURES	Location/Qualifiers
source	1. .1964

```

/organism="Perilla frutescens"
/cultivar="Okdong"

```

```

nena
1 1964
/db_xref="taxon:48386"
/curcivus = OROONGY

```

```

1. .1304
/ gene="Dgat1"
50 1673
cns

```

CDS
69..1073
/gene="Dgat1"
/locus="Dgat1"

```

/Note="DGA1"
/codon_start=1

```

```

/product="diacylglycerol acyltransferase"
/protein_id="AAG23696.1"

```

```

/db_xref="GI:10803053"
/translation="MAILSP EILDTSSADNGAAHHTTLRRRQSARSVPPLDSDS

```

NSLEAESAINDSVNRNDANLIENLRGGAVESENEKEQSYGKBECAKVKENGFTSNG
 NSLEAESAINDSVNRNDANLIENLRGGAVESENEKEQSYGKBECAKVKENGFTSNG
 GTDVMVKTFRPAAPAHRRKNKESPLSSDAIFKQSHAGLFNLCIWLVAVNSRLIEN

LMKYGWLTKSGFWFSSTSLRDWPLLMCCLSLPVFALASFLVKLVKLYNIPETWAVEL
HYTITTVLEFPVVILRCDSAVLSGVTLMFACTVWLKVSFAHTNDLRLAKSLD

KWEAMSRYNLDYAYDVSPKSLAYFWVAPTLCYQPSYPTACTRKGWVVRQLIKLVIF
TGLMGFIIEQYINPIVQNSQHLKGNLLYATERVKLVSVPNLVYVLCWFYCFFFHLWLN

ILAE L L C F G D R E F Y K D W N A R T V E Y W R M N N P V H K M V R H Y I Y C P C L Q N G I P K I V A V L
I A F L V S A I F H E L C V A P C Q I F K F W A P S G I M L Q V P L V I V T N Y L O E K F K S M S V G N M M F W C

```

BASE COUNT      522 a  395 c  461 g  586 t
FTCIFGQPCVCLYYHDLNMRKASAR"

```

ORIGIN

Query Match 41.7%; Score 810.2; DB 8; Length 1964;
Best Local Similarity 76.8%; Pred. No. 1.3e-176;

Matches 989; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 243 AAGACACTGATTCTTCGCTCCTCAAATTCGGCTACGCTCCTTCCGTCCCGCGCTACCGCA 302

Db 373 ACGGAAC TGATGTTATGGCCGTCAAATTCACATTCAGGCCGGGGGGCCTGCTACCGCA 432

QY 303 AAGTGAAGGAAAGTCCGCTCAGCTCCGACACTATTTTCCGTCAGAGTCACGGGGCCTCT 362

Db 433 AAAATAGGAGAGTCTCTTAGCTCCGAGCCATCTTCAACAGAGCCATGCAGGCTCT 492

Qy 363 TCAACCTTTGTATAGTAGTCCTTGTTGCTGTGAATAGCCGACTCATCATTTGAGAATTAA 422

Db 493 TCAACCTTTGTATAGTGGTGCTTGTTGCTGTAATAAGCAGACTAATATTGAGAAATTAA 552

QY 423 TCAAGTATGGTTGGTTGATCAAAATCTGGCTTTTGGTTTAGTTCAAAGTCAATTGAGAGACT 482

Db 553 TGAAGTATGGGTGGCTGATCAAAATCAGGATTTGGTTTAGTTCAACATCGCTTAGGGATT 612

Qy 483 GGCCCCCTTTTCATGTGTGTCTCTTTCTCTCTCTGTGGTATTTCCCTTCGCTGCCTTTATAGTG 542

Db 613 GGGCACTGCTAATGCTGTGTCTTACTCTTCCAGTTTTTGCACTCGCTTCATTTCTGTGG 672

QY 543 AGAAGTTGGCCACAACGGGAAGTGTATACCCGAACCACTTGTTGTTGTACTTCAATATAATCA 602

Db 673 AGAAGTTGGTGAACCTAAATTATAPACCTGAGTGGGTCCGACCTCTTCTTCATGTTACAA 732

QV	603	TITACCTCAACTTCGGCTTTTCTATCCAGTTTTAGTATTCTCAGGTGCAATTCGCTTTTG 662
----	-----	--

[illegible]

20	53	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	6
----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---

[illegible][illegible][illegible]

DB 853 CACATACAACTATGATTGAGAGTACTTGC AAAATCACITTGATAAGTGGGAGCTATGT 912

$\frac{1}{2} \leq \frac{1}{2} \leq \frac{1}{2}$

Qy	783	TCGATACTCTGAAACATGGAGCTATCCCTTTACAACGTAAGCTTCAAGAGCTTGGCATATTTCC	842
Db	913	CCAGGTACTGGAACTCCGACTACGCTTATGATGTAAAGCTTTAAGAGCTCGGCATACTTCA	972
Qy	843	TGGTTGCCCTACATATATGTTACCGCAAGCTATCCCTCGCACACCTTATATTTCGAAGG	902
Db	973	TGGTTGCTCTAGATTGTGTTACAGCCAAGCTACCTCTGGACAGCTTGTGCATTCGGAAAG	1032
Qy	903	GTTGGTTGTTTTCGCCAACTTGTCCAAGCTCATATATTTTACAGAGCTTATGGGATTTATAA	962
Db	1033	GTTGGGTGGTAAGCAACTAATTTAAGCTGGTAAATATTCACAGAGCTCATGGGATTTATTA	1092
Qy	963	TAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCCTCTCAAGGGAACCTTC	1022
Db	1093	TAGAACAGTACATAAAACCGGATCGTTCAAAATTCACAACCTCTCTGAAAGGAAACCTTT	1152
Qy	1023	TTTAGGCCACCGAGAGAGTCTGGAAGCTTCTGTTTCCAAAATTTATATGTGTGGCTCTGCA	1082
Db	1153	TATATGCCATTGAGAGGGCTCTTGAAGCTTCTGTTTCCAAAATTTATATGTGTGGCTCTGCA	1212
Qy	1083	TGTTCTATTCTCTTTTCCACCTTTGGTAAATATATCCTGCGACAGAGCTTCTCGATTGGTG	1142
Db	1213	TGTTTATTCTGTTTTTCCACCTCTGGCTAAATATACTTGTCTGGAAGCTCTGTGCTTTGGGG	1272
Qy	1143	ATCGTGAATTTCTACAAGGATTTGGTGGAAATGCCAAAACCTGTCCGAAGATTTATTTGGAGGATGT	1202
Db	1273	ACCGTGAGTTTTATAAGGATTTGGTGGCAATGCGAGGACAGTGGAGGAGTACTGGAGAAATGT	1332
Qy	1203	GGAAATGCGCTGTTTCAAAAATGGATGATCCGCCACCTATATATTTTCCATGTTTTAGGCACG	1262
Db	1333	GGAAATGCGCTGTCCATAAATGGATGGTTTCCGCGCATATACTGTCCATGCTTTACAAAATG	1392
Qy	1263	GTCCTACCAAAAGGCTGCTGCTCTTTTAAATGGCTTCTCTGCTTTTCTGCTTATTTCCATGAGC	1322
Db	1393	GAATACCAAGATAGTGGCAGTTTTGATCGCGGTTTCTTGTGTCTCGGATTTTTTCATGAGC	1452
Qy	1323	TGTGCAATGCTGTTCCTTGCCACATATTCAAAGTTTGTGGGCTTTCCGGTGGAAATATGTTTC	1382
Db	1453	TGTGCGTTTGCAGTCCCTTGCCAAATATTCAAGTTTTTGGGCGTTCTTCGGGTATCATGCTTC	1512
Qy	1383	AGTTTCTTTGGTCTTGATCAGTCAATATTTCTGCAAAATTAATTCAGAAACTCAATGGTTG	1442
Db	1513	AGTTTCTCTCGTAATCGTGACTAATTTACTTCCAGAAAAGTTTCAAAAACCTCAATGGTGG	1572
Qy	1443	GAATATGATTTTTTGGTTTCATATTCAGTATCCTTTGGTCAACCTATGTGTGCTACTGCTAT	1502
Db	1573	GCAATATGATGTTCTCGGTGGTCTTCTGTATCTTTGGTCAACCTATGTGTGTGTTGCTGT	1632
Qy	1503	ACTACCATGACTTGTATGAATAGAAAG	1529
Db	1633	ACTACCACGACTTGTATGAATTCGAAAG	1659

RESULT 2	AY084052	LOCUS	2090 bp	mRNA	linear	PLN 08-APR-2002
DEFINITION						
Tropaeolum majus putative diacylglycerol acyltransferase mRNA, complete cds.						

COMPLETE COPY:	
ACCESSION	AY084052
VERSION	AY084052.1
KEYWORDS	GI:20086323
SOURCE	.
ORGANISM	Tropaeolum majus.
	Tropaeolum majus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Tropaeolaceae; Tropaeolum.
(1 (bases 1 to 2090))

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE

Mietkiewska,E., Pedersen,K., Katavic,V. and Taylor,D.C.
Characterization of a putative diacylglycerol acyltransferase mRNA from Tropaeolum majus embryo
Unpublished
2 (bases 1 to 2090)

Db	1058	GCCTAGCTATCTCCGTTCTGCTGATATCCGCAAGGGTGGTGTTCGTCAATTTGTCAA	1117
QY	928	GCTGATAATATTTACAGGAGCTATGGGATTTATAATAGACCAATATATTAATCCCATAGT	987
Db	1118	ACTAATAGTTTTTCATAGGACTCATGGGTTTCATATAGACAATATATTAATCCTATCGT	1177
QY	988	ACAAATTCACAGATCCTCTCAAGGAAACCTTTTACGGCACCGAGAGAGTTCTGAA	1047
Db	1178	TCGAAATTCGCAACACCCATTGAAGAGAGATTTTATATGCAATAGAAAGAGTTTGAA	1237
QY	1048	GCCTTCTGTTCGCAAAATTTATATGTGGGCTCGCATCTTCTATTTGCTTTTCCACCTTTG	1107
Db	1238	GCCTTTCAGTTTCCAAATCTATATGTTGGCTTTGCATGTTCTACTCTTTTCCACCTCTG	1297
QY	1108	GTTAAATATCCTGGCAGAGCTTCTTCGATTTGGTGATCGTGAATCTTACAAAGGATTCGTG	1167
Db	1298	GTTGACATACCTGGCTGAGCTTCTTCGCTTTGGTGATCGTGAATCTTACAAAGATTCGTG	1357
QY	1168	GAATGCCAAACTGTGCAAGATTTATTGGAGGATGTGGAAATATGCTGTTTCCAAATGGAT	1227
Db	1358	GAATGCCAAACTGTTGCGGAGTATTGGAATAATGGAATATGCTGTTTCCATAGATGGAT	1417
QY	1228	GATCCGCCACCTATATTTTCCATGTTTAAAGCGAGGCTTACCAAGGCTGCTGCTCTTTT	1287
Db	1418	GTTTCTGCTCATCTATATTTTCCCTGTTTGGAGAAATGGGATACCCAAAGGAGTGCCATTAT	1477
QY	1288	AATTGCTTCTCGTGTCTCTTTATTTCCATGAGCTGTGCATTCGCTGCTTCCCTGCCACAT	1347
Db	1478	TATCGGGTCTTATAGTTTCTGGTGTTCCTTCCATGAGCTCTGCATTCGCTGCTTCCCTGCCACGT	1537
QY	1348	ATTCAAGTTTGGGCTTTCGGTGGAAATATGTTTTCAGGTTTCCCTTGGTCTTGATCACTAA	1407
Db	1538	ATTCAAGTTTATGGGCTTTTATAGGCATTTATGTTTTCAGGTTTCCCTTGGTATGATTACGAA	1597
QY	1408	TTATCTGCAAAATTAATTCAGAACTCAATGTTGGGAAATATGATTTTGGTTGTTTCATATT	1467
Db	1598	TTATCTACAGAAAGTTCTAGTAATTTCTATGGTGGGCAATATGATCTTCTGGTTTCATCTT	1657
QY	1468	CAGTATCTTGGTCAACCTATGTTGTGACGTCTATCTACCATGACATGATGAATAGGAA	1527
Db	1658	CTGCATCTTGGGCAACCTATGTTGTCCTTCTATATTCAGTACCTGATTAATCTTAA	1717
QY	1528	AGGCAA	1533
Db	1718	GGAAAA	1723
RESULT 3			
AF129003			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

FEATURES
source

Location/Qualifiers
1..2099
/organism="Nicotiana tabacum"
/cultivar="xanthi SH6"
/db_xref="taxon:4097"
/tissue_type="calli derived from leaf protoplasts"
1..2099
/gene="DAGAT"
208..1806
/gene="DAGAT"
/function="catalyzes the acylation of diacylglycerol by oleoyl-CoA to give triacylglycerol"
/note="similar to acyl-CoA cholesterol acyltransferase"
/codon_start=1
/product="diacylglycerol acyl-CoA acyltransferase"
/protein_id="AAF19345.1"
/db_xref="GI:6625653"
/translation="MVMLPESEVEMTTTTTSIENLNSDLNHSVRRRSGNSFEAA
SAINSSDNMSDEDRDGGSGAGLETWNERSKSGESDVIRKEDDRNDVANGEEK
STETTPPFAYRASAPAHRRIKESPLSDAIFKQSHAGLIVLCVWVLIIVANERLII
ENLMYGLIRAGFWFSKSLRDWPLMCCLSIQLPLAFLVLEKLAQORHLTERAVV
TLHITTAALYPLVILGCDSDAFLGVILMLVACILVWMLKLSYAHNDHMRLOAKS
TDKDETSDGSDYDVSFKSLAYENVAPTLCSYHPHTPCIRKSWARQFKLVIPTG
LMGFIQYINPIVONSHPLKGNLLIAERVLKSLVNLVWLCMTFPPHMLNIL
AELLCFGDREYIKDWMNAKTIDERYWNNMPVHKWVRHYFPCLRNGIPKGVAILIA
FLVSAVFELCIAYPCRLFKWAFMGIMFQVPLVILTNFLQNKFSQSVGNMFWCFE
CILQPMCVLLIYHDVNMKRSAR"
BASE COUNT 543 a 405 c 485 g 666 t
ORIGIN

Query Match 38.3%; Score 744.6; DB 8; Length 2099;
Best Local Similarity 73.9%; Pred. NO. 1.9e-161;
Matches 964; Conservative 0; Mismatches 329; Indels 12; Gaps 1;

QY 228 AGCAAAACGAAACACACACTGATTTTCGCTCCTCAAAATTCGCTACCGTCTCTCG 287
DB 503 AGGAAGCAAAATCAACGGAACACACAGCGCCGCTTTAAATTTTCTTACAGGCGCTCG 562
QY 288 TCCCGCTCACCGCAAGTGAAGGAAGTCCGCTCAGCTCCGACACTATTTTCCGCTCAGA 347
DB 563 CACCAAGTCCCGCGGGAATCAAGGAGAGTCTCTCAGCTCCGACGCCATTTTCAACACAGA 622
QY 348 GTCACGCGGCGCTCTTCAACCTTTCTAGTAGTCTCTTGTGCTGTGAATAGCCGACTCA 407
DB 623 GTCACGCGAGGCGTCTCAATCTCTGTGTGTGTCTGCTGATTTTCAACAGCAGGCTGA 682
QY 408 TCATTGAGAAATTAATGAAGTATGCTTGTGATCAATCTGCTTTTGTGTTTACTTCAA 467
DB 683 TTATCGAAGACTGTAGTGAAGTATGCGCTTTAAATGAGGCGTGTGTTTGTGTTTACTCGA 742
QY 468 AGTCATTGAGAGACTGGCCCTTTTCATGTGTGTCTTTCTCTTGTGTATTTCTCTTCG 527
DB 743 AGTCGTTGAGGATTTGGCGCTTCTTAATGTCTGTCTCAGTCTCCAAATTTTGGCGCTCG 802
QY 528 CTGCGCTTTATGAGTGGAGAGTTGGCACAAAGGAAGTGTATACCCGACAGTGTGTTGTTG 587
DB 803 CTGCTTTTCTTGTGAGAGTGTGGCACAGAGGCAATTTGACTGAGCGTGGGTGTTA 862
QY 588 TACTTCATATATCATACCTCACTTCTTCTTCTATCCAGTTTCTTGTATTTCTCAGT 647
DB 863 CTCTTCACATACATACACAGCTGCCATTTGTATCCAGTTCTGTCATCTTGTGTTGTTG 922
QY 648 GTGATTCGCTTTTGTATCAGGTGTACGTTAAATGCTGTTTCTTGTGTTGTTGTTGTTAA 707
DB 923 GTGATTCGCTTTTCTTGTGTTGTTCAATATGATGCTGTTGTTGTTGTTGTTGTTGTTA 982
QY 708 AATTCGCTCTTTATGACATACAAACTATGATATGAGAGCACTTACCAAAATAGTTGAAA 767
DB 983 AGCTGGTTTCTTACGCATACAAATCATGATATGAGAGCACTGCAAAAGTCTACGGACA 1042
QY 768 AGGGAAGAGCACTGCTCGATCTCTGAACATGGAGTATCTTACAACTAGCTTCAAGA 827
DB 1043 AGGATGAA-----ACTTCAGATGGGGATTTCTCTTATGATGTTAGCTTCAAGA 1090

QY 828 GCTTGGCATATTTCTCGTGGTCCCTCATATTTTACAGCAAGCTATCTCCACAC 887
DB 1091 GTTGGCTTACTTATGTTGGTGGCCAAACATTTATGTTATCAGCTTACCTCCACACTC 1150
QY 888 CTATATTTCCGAAAGGTTGGTTTCCCAACTTGTCAAGCTGATAATATTTACAGGAG 947
DB 1151 CATGATTTCCGAAAGGTTGGTGGCAGCCCAATTCATCAAGCTGCTAATATTTACAGCAT 1210
QY 948 TTATGGGATTTATATAGACCAATATATTAATCCCATAGTACAAATTTACAGCATCTC 1007
DB 1211 TGATGGGATTTATCATAGAACAGTATTAACCCAAATTTGTGCAAAACTCACAACATCCTT 1270
QY 1008 TCAAGGGAACCTTTCTTACGCCAGCAGAGAGTTTCTCAAGCTTCTGTTCACAAATTTAT 1067
DB 1271 TGAAGGGAACCTTTTATATGCCATCGAGAGGGTATTTGAGCTTTCGGTTCGCAATTTAT 1330
QY 1068 ATGTGTGCTCTGCATGTTCTATTTCTTATTTTCCACCTTTTGGTTTAAATATCTCGCAGAGC 1127
DB 1331 ATGTCTGCTCTGCATGTTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1390
QY 1128 TTCITCGATTTGGTGCATCGTGAATTTCTACAGGATTTGGTGAATCCCAAACTGTCGAAG 1187
DB 1391 TACTATGTTTGGTGCATCGTGAATTTCTACAGGATTTGGTGAATCCCAAACTGTCGAAG 1450
QY 1188 ATTATGGAGGATGTGGAATATGCTGTTTCAAAATGATGATCGCCACCTATATATTTTC 1247
DB 1451 AGTACTGGAGGATGTGGAATATGCTGTTTCTATAGTGGATGTTGCTGCTACATTTATTTCC 1510
QY 1248 CATGTTTAAAGCAGCGTCTACAAAGGCTGTGCTCTTTTAAATGGCTTCTCGTGTCTG 1307
DB 1511 CTGCTTAAAGAAACGAATTTCCAAAGGGGTCGCAATACTGATTTCTTCTTCTTCTTCTG 1570
QY 1308 CTTTATTCATGAGCTGTGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1367
DB 1571 CTGTTTCCAGAGCTGTGATTTGCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1630
QY 1368 GTGAAATTTATTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1427
DB 1631 TGGGAATTTATTTCTCAGGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1690
QY 1428 GAACTCAATGGTGGAAATATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1487
DB 1691 AAAGTCGATGCTGGCAATATGATGTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTG 1750
QY 1488 TGTGTGACTGCTATCTACTACCATGATGATCAATAGGAAGCA 1532
DB 1751 TGTGTGCTCTTCTGTTATTTACCCAGGATGTGATCAATAGAAAGCA 1795
RESULT 4
ATH238008 1904 bp mRNA linear PLN 18-JUN-1999
LOCUS Arabidopsis thaliana mRNA for diacylglycerol acyltransferase.
DEFINITION Arabidopsis thaliana mRNA for diacylglycerol acyltransferase.
ACCESSION AJ238008
VERSION AJ238008.1 GI:5123717
Dated gene: diacylglycerol acyltransferase.
KEYWORDS Arabidopsis thaliana.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1904)
Zou J., Wei Y., Jaki C., Selvaraj, G. and Taylor, D.C.
The Arabidopsis thaliana TAG1 gene encodes for a diacylglycerol
acyltransferase
Unpublished
REFERENCE 2 (bases 1 to 1904)
Zou J.
Direct Submission
Submitted (06-APR-1999) Zou J., Plant Biotechnology Institute,
National Research Council of Canada, S7H 0W9, Saskatchewan, CANADA
Location/Qualifiers
FEATURES

```
source 1. .1904
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
gene 139. .1701
/genes="Dgat"
CDS 139. .1701
/genes="Dgat"
/function="lipid synthesis"
/codon_start=1
/product="diacylglycerol acyltransferase"
/protein_id="CA45373.1"
/db_xref="GI:5123718"
/translation="MAILDSAGVTVTENGCGEFVDLRLRRKSRSDSSNGLLLSGS
DNPSDDVGAADYRDRISSVNDQAGTANLADNNGGDDNGGGRGGGRGNAD
AFTTYRPSDVAARESPLESDDAIFKQSHAGLFNLVYVLIANSLRIENLMRYG
LIRDFWSSRLSDWPLFMCCISLIFPLAAFTVEKLVQYISRPVVFILHIITM
TEVLXPVYTLRCDASFLSGVTMLLTCLVWLKLSVYAHYSYDIISLANADKANPEV
SYVSLKSLAYPMVAPTLCYOPSPRSACIRKGMAROPAKIVFTGPMGFILQYIN
PIVRNKHPLKGLDILYATERVILKSVPNLYVWLMPYCFPHLWILAILLAEFLCGDBREF
YKDMNNAKSGDYWRMNNPVHKWVRIYFPLRSKIPKTLAILLILAEFLVAFVHELIC
IAVPCRLFKLAFGLIMFQVPLVFTTNVLOERFSGTVGNMIFFCIFGQPMCVLLY
YHDLNMRKGSMS"
BASE COUNT 458 a 399 c 431 g 616 t
ORIGIN

Query Match 37.1%; Score 719.8; DB 8; Length 1904;
Best Local Similarity 73.2%; Pred. No. 1e-155;
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;

QY 241 ACAAGACACTGATTTCTCGCTCCTCAAAATTCGCTACCGTCCTCGCTCCGCTCACCG 300
DB 423 AGAAGGAAGGAAGACCGCGATGCTAGTTAGTATCGACCGCTCGGTCACGCTCATCG 482
QY 301 CAAAGTGAAGGAAGTCCGCTCAGCTCCGACACTATTTTCCGTCAGAGTCACCGCGGCT 360
DB 483 GAGGCGGAGAGAGTCCACTTAGCTAGCTCCGACGCAATCTTCAACACAGAGCCATCGCGATT 542
QY 361 CTTCAACCTTTGTATAGTAGTCTTGTGCTGTGAATAGCCGACTCATCTTACAGATTT 420
DB 543 ATTCAACCTCTGTGTAGTAGTCTTGTGCTGTGAATAGCCGACTCATCTTACAGATTT 602
QY 421 AATCAAGTAGTGTGTGTATCAAACTTGGCTTTTGGTTTGTATCAAGTCAATTCAGAGA 480
DB 603 TATGAAGTAGTGTGTGTATCAAGAGGATTTCTGGTTTGTATCAAGATCGCTGCGAGA 662
QY 481 CTGCGCCCTTTTCATGTTGTTCTTCTCTGTGTTATTTCCCTTCGCTGCCCTTTATAGT 540
DB 663 TTGCGCGCTTTTCATGTTGTTATATCCCTTTTCGATCTTTTCCCTTGGCTGCCCTTACGGT 722
QY 541 GGAGAAGTTGGCACACGGAAGTGTATACCCGAACAGTTGTTGTTGTTACTTTCATATAAT 600
DB 723 TGAGAAATTGGTACTTCAGAAATACATATCAGAACCTGTGTGTCATCTTTCTTTCATATTAT 782
QY 601 CATTACTCAACTTCGCTTTTCTATCCAGTTTGTATTTATTTCTCAGGTGTGATTCGCTTT 660
DB 783 TATCACCATGACAGAGTTTGTATCCAGTTTACGTCACCCCTAAGGTGTGATTCGCTTT 842
QY 661 TGTATCAGGTGTGCACGTTAATGCTGTTTCTTGTGTTATGTTTAAATFTGGTGCTTTA 720
DB 843 TTTTATCAGGTGTGCACGTTTGTATGCTCTCCTCACTTGCATTTGTGGCTAAAGTTGTTCTTTA 902
QY 721 TGCACATACAAACTATGATGAGACGCTTACCAAAATTTAGTTTGAAGGAGGAGACACT 780
DB 903 TGTCTACTAGTATGACATAAGATCCCTTAGCCAAATGACAGCTGATAAGG----- 952
QY 781 GCTCGATATCTCAACATGGAGTATCCTTACACAGTAAAGTTCACAGGCTTGGCATATTTT 840
DB 953 -----CCAACTCGAAGTCTCTACTACGTTAGTTTGAAGAGCTTGGCATATTT 1001
QY 841 CTTGGTTGCCCTTACATTTATGTTTACCAGGCAAGCTATCTCTCGCACACCTTTATTTGAAA 900
DB 1002 CATGGTCGCTCCACATTTGTTTATCAGCAAGTTTATCCACGTTCTTCATGTTATACGGAA 1061
```

```
QY 901 GGGTTGGTTGTTTCCGCCAACTTGTCAAGCTGATAATATTTACAGGAGTGTATGGATTAT 960
DB 1062 GGGTTGGTTGGCTGCTGCAATTTGCAAACTGGTGCATATTCACCGGATTCATGGATTAT 1121
QY 961 AATAGACCAATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020
DB 1122 AATAGAACCAATATTAATCCCATAGTTCAGGAACCTCAAGCATCTCTTGAAGGCGATCT 1181
QY 1021 TCTTTACGCCACCGAGAGAGTCTTGAAGCTTTTCCACCTTTGTTTCCAAAATTTATATGTGTGGCTCTG 1080
DB 1182 TCTATATGCTATTGAAAGAGTGTGAAGCTTTTCAGTTCCCAATTTATATGTGTGGCTCTG 1241
QY 1081 CATGTTCTATGCTTTTCCACCTTTGTTTAAATATATCTTGGCAGAGCTTCTTCGATTGG 1140
DB 1242 CATGTTCTATGCTTTTCCACCTTTGTTTAAATATATTCGACATATTCCTCCGCTTGGCAG 1301
QY 1141 TGATCTGTAATCTCAAGAGATTGTGGAATGCCAAACTGTCCAAAGTGTGAAGATTATTTGGAGAT 1200
DB 1302 GGATCTGTAATCTCAAGAGATTGTGGAATGCCAAAGTGTGGAGATTACTTGGAGAAT 1361
QY 1201 GTGGAATATGCCTGTTTCACAAAATGATATCCGCCACCTATATTTTCCATGTTTAAAGGCA 1260
DB 1362 GTGGAATATGCCTGTTTCAATAATGATGTTTCGACATATATATCTCCCGTGTGGCAG 1421
QY 1261 CGGCTTACCAAGGCTGCTCTCTTTTAAATGCTTCCCTGGTTTCTGCTTTATTCATGA 1320
DB 1422 CAGATACCAAGACACTCGCCATTTATCATTTGCTTCTCCTAGTCTCTGCAGCTCTTTTATGA 1481
QY 1321 GCTGTCAATGCTGTTTCCCTGTCACATATTTCAAGTTGTGGCTTTCGGTGGATTATGTT 1380
DB 1482 GCTATCGATCGAGTTCCTTGTGCTCTCTTCAAGCTATGGCTTTTCTTGGGATTATGTT 1541
QY 1381 TCAGGTTCTTTGGTCTTGTATCACTAATATCTGCAAAATATAATTCAGAACTCAATGTT 1440
DB 1542 TCAGGTCCTTTGGTCTTTCATCAAACTATCTACAGGAAGGTTTGG---CTCAAGCT 1598
QY 1441 TGGAAATATGATTTTGGTTCATATTCAGTATCTTGTTCACACCTATGTGTGCTACTGCT 1500
DB 1599 GGGGAACATGATCTTCTGTTTCTCTCTGATTTTCGACATTTTCGACAAACCGATGTGTGCTTCT 1658
QY 1501 ATACTACCAGTCTGATGAATAGGAAGCAAACTTGACTGAAGCTACGG 1551
DB 1659 TTATTACCAGCTGATGNAACCGAAAGGATCGATGTCATGAACAACATG 1709

RESULT 5
LOCUS AX090345 1942 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 38 from Patent WO0116308.
ACCESSION AX090345
VERSION AX090345.1 GI:13444207
KEYWORDS thale cress.
SOURCE ORGANISM
Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1942)
AUTHORS Lasser M. and van Eenennaam A.
TITLE Plant sterol acyltransferases
JOURNAL Patent: WO 0116308-A 38 08-MAR-2001;
MONSANTO COMPANY (US)
FEATURES
Source
1. .1942
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 461 a 421 c 425 g 635 t
ORIGIN

Query Match 37.1%; Score 719.8; DB 6; Length 1942;
Best Local Similarity 73.2%; Pred. No. 1e-155;
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;
```

QY 241 ACAAGACACTGATTTCTCGTCCCTCAAAATTCGGCTACCGTTCCTCCGTCCTCCGCTCACCG 300
Db 521 AGAAGGAGAGGAAACGCGGATGCTACGTTTACGATATCGACGCGGTTCAGCTCATCG 580
QY 301 CAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTAATTTCCGTCAGAGTCCAGCGGGCT 360
Db 581 GAGGCGGAGAGAGTCCACITTAGCTCCGACGCAATCTTCAACAGAGCATCGCGGAT 640
QY 361 CWTCAACCTTTGTATAGTAGTCTTGTGCTGTGAATACCGGACTCATCATTTGAGAATTT 420
Db 641 ATTCAACCTCTGTGTAGTAGTCTTGTGCTGTGAATACCGGACTCATCATCGGAAATCT 700
QY 421 ATAGAGTATGTTGGTGTGATCAAAATTCGGCTTTTGGTTTATAGTCAAAAGTCAATGAGAGA 480
Db 701 TATGAAGTATGGTTGCTATCAGAACGATTTCTGTTTATAGTTCGAAGTCCGTCGAGA 760
QY 481 CTGGGCCCTTTTCATGTGTGCTTCTCTGTGCTATTTCTCTTCGCTGCTTTATAGT 540
Db 761 TTGGCGGCTTTTCATGTGTGATATCCCTTTCGATCTTCTTGGCTGCTTTACGGT 820
QY 541 GGAGAAAGTTGGCACAACGGAAGTGTATACCGGAACAGTGTGTGTACTTCATATAAT 600
Db 821 TCAGAAATTTGGTACTTCAGAAATACATATCAGAACCTGTGTCTATCTTTCATATAT 880
QY 601 CATTAACCTCAACTTCGCTTTTCTATCCAGTTTATAGTATTCAGGTGTGATCTGCTTT 660
Db 881 TATCACCATGACAGAGTTTGTATCCAGTTTATACCTACCCCTAAGGTGTGATCTGCTTT 940
QY 661 TGTATCAGGTGTCAGTTAAAGTCTGTTTCTGTCTGTGTATGGTTAAATTTGGTCTCTTA 720
Db 941 TTTATCAGGTGTCATTTGATGCTCCTCACITTCGATGTGTGGCTTAAAGTTGGTTCTTA 1000
QY 721 TGCACATACAAACTATGATATGACAGCACTTTACCAAAATTTAGTTGAAAGGGAGAGCACT 780
Db 1001 TGCTCATACTAGTATGACATAAGATCCCTAGCCCAATGCAGCTGATAAG ----- 1050
QY 781 GCTGATACTCTGAACATGGACTATCTTTACACCTAAGCTTCAAGAGCTTGGCATATTT 840
Db 1051 -----CCAACTCTGAAGTCTCCTACTACCTTGAAGAGCTTGGCATATTT 1099
QY 841 COTGCTTGCCCTACATTTATGTTACACGCAAGCTATCTCGCACACTTATATTCGAAA 900
Db 1100 CATGTCGCTCCACATTTGTGTATCAGCCAAAGTTATCCAGCTTCGATGTATACGGA 1159
QY 901 GGGTTGGTTGTTTCCCAACTTGTCAAGCTGATATAATTTTACAGAGTATGGGATTTAT 960
Db 1160 GGGTTGGTGGCTCTCAATTTGCAAACTGTGTATATTCACCGGATTCATGGGATTTAT 1219
QY 961 AATAGAACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020
Db 1220 AATAGAACCAATATATAATCCCTATTTGTGAGAACCTCAAAAGCATCTCTCTGCTCG 1279
QY 1021 TCTTTACGCCACCGAGAGTCTCAAGCTTCTGTTCCAAATTTATATGTTGGCTCTG 1080
Db 1280 TCTATATGCTATTGAAGAGGTTGAAAGCTTTCAGTTCCAAATTTATATGTTGGCTCTG 1339
QY 1081 CATGTTCTATTGCTTTTCCACCTTTGGTTAAATATCTCGCAGAGCTTCTTCGATTTGG 1140
Db 1340 CATGTTCTACTGCTTCTTCCACCTTTGGTTAAACATATTTGGCAGAGCTTCTCTGCTCG 1399
QY 1141 TGATCGTGAATCTACAGAGTGTGTGAATGCCCAGGCTTCGAGAGTATTTGGAGAT 1200
Db 1400 GATCGTGAATCTACAAAGATTTGGTGAATGCAAAAAGTTGGGAGATTTACTGGAGAT 1459
QY 1201 GTGGAATATGCTGTTCAAATGATGATCATCGCCACCTATATTTTCCATGTTTAAAGCA 1260
Db 1460 GTGGAATATGCTGTTCATTAATGAATGGTTGAGATATATCTTCCGCTGCTTCCGAG 1519
QY 1261 CGGTCTACCAAGGCTGCTGCTTTTAAATTCCTTCTGTTTCTGCTTTATTCATGA 1320
Db 1520 CAAGATACCAAGACACTCGGCATATATCATTTGCTTCTAGTCTCTGAGCTTTTCATGA 1579
QY 1321 GCTGTGCATTTGCTTCTTCCCATATATTCAGTTTGTGGGCTTTCGGTGGAAATATGTT 1380

Db 1580 GCTATGATCGAGTTCCTTGTGCTCTCTTCAAGCTATATGGCTTTCTTGGGATATGTT 1639
QY 1381 TCAGTTCCTTTGGTGTGATCACTAATTTATCTGCAAAATAAATTCAGAACTCAATGTT 1440
Db 1640 TCAGTGTCTTTGGTGTTCATCACAACACTATCTACAGGAAGGTTTGG---CTCAACGGT 1696
QY 1441 TGGAAATATGATTTTGGTTCATATTCAGTATCCCTTGGTCAACCTATGTGTACTGCT 1500
Db 1697 GGGGAACATGATCTTCTGTTTCATCTTCTGATTTTCGGACAACCGATGTGTGCTTCT 1756
QY 1501 ATACTACCATGACTGTGATGAATAGGAAGGCAAACTTCACTGAGACTAGG 1551
Db 1757 TTATTACCACGACTGATGAACCGAAAGGATCGATGTCATGAACAACATG 1807
RESULT 6
AX090349 AX090349 1942 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 42 from Patent WO0116308.
DEFINITION AX090349
ACCESSION AX090349
VERSION AX090349.1 GI:13444210
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 1942);
AUTHORS Luesner, M. and van Eenennaam, A.
TITLE Plant sterol acyltransferases
JOURNAL Patent: WO 0116308-A 42 08-MAR-2001;
MONSANTO COMPANY (US)
FEATURES
Location/Qualifiers
Source 1..1942
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 461 a 421 c 425 g 635 t
ORIGIN
Query Match 37.1%; Score 719.8; DB 6; Length 1942;
Best Local Similarity 73.2%; Pred. No. le-155;
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;
QY 241 ACAAGACACTGATTTCTCCGTCCTCAAAATTCGGCTACCGTTCCTCCGCTCACCG 300
Db 521 AGAAGGAGAGGAAACGCGGATGCTACGTTTACGATATCGACGCGGTTCAGCTCATCG 580
QY 301 CAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTAATTTTCCGTCAGAGTCCAGCGGGCT 360
Db 581 GAGGCGGAGAGAGTCCCACTTAGCTCCGACGCAATCTTCAACAGAGCCATGCCGGATT 640
QY 361 CTTCAACCTTTGTATAGTAGTCTTGTGCTGTGAATAGCGAGCTCATCATTTGAGAATTT 420
Db 641 ATTCAACCTCTGTGTAGTAGTCTTATTGCTGTAAACAGTAGACTCATCTCGAAAATCT 700
QY 421 AATGAAGTATGTTGGTTGATCAAAATCTGGCTTTTGGTTTATAGTCAAAAGTCAATGAGAGA 480
Db 701 TATGAAGTATGTTGGTTGATCAGAACGATTTCTGTTTATAGTTCAGATCGCTCGAGA 760
QY 481 CTGGCCCCCTTTTCATGTGTGCTTCTTCTTGTGTTGTTATTCCTTTCGCTGCTTTATAGT 540
Db 761 TTGGCCGCTTTTCATGTGTGTTGTATATCCCTTTCGATCTTCTTGGCTGCTTTACGGT 820
QY 541 GGAGAAAGTTGGCACAACGGAAGTGTATACCGGAACAGTGTGTGTACTTCATATAAT 600
Db 821 TCAGAAATTTGGTACTTCAGAAATACATATCAGAACCTGTGTCTATCTTTCATATAT 880
QY 601 CATTAACCTCAACTTCGCTTTTCTATCCAGTTTATAGTATTCAGGTGTGATCTGCTTT 660
Db 881 TATCACCATGACAGAGTTTGTATCCAGTTTATACCTACCCCTAAGGTGTGATCTGCTTT 940
QY 661 TGTATCAGGTGTCAGTTAAAGTCTGTTTCTGTCTGTGTATGGTTAAATTTGGTCTCTTA 720

Db 995 TCTCATACTAGTATGACATGAAGATCCCTAGCCAAATCAGCTGATAAGG----- 1044
Qy 781 GCTGATACCTCTGAACATGAGACTCTCTTACAACTGAAGCTTCAAGAGCTTGGCATATTT 840
Db 1045 -----CCATCTCTGAAGTCTCTCTACTACTAGCTTGAAGAGCTTGGCATATTT 1093
Qy 841 CCGTGTGCTCCCTACATATGATTTACCAAGCTATCTCGCACACCTTATATTCGAAA 900
Db 1094 CATGTCCTCCACATATGTTATCATCAGCAAGTATCCAGCTTCTGCATGTATACGGA 1153
Qy 901 GGGTGGTGTGTTCCCAACTGTCAAGCTGATATATTTACAGAGTATGGGATTTAT 960
Db 1154 GGGTGGTGGTCTGCTCAATTTGCAAACTGGTCTATATTCACCGGATTCATGGGATTTAT 1213
Qy 961 AATGACCAATATATATATCCATAGTACAAATTCACAGCATCTCTCAAGGAAACCT 1020
Db 1214 AATGACCAATATATATATCTTGTTCAGCAACTCAAGCATCTCTTGAAGGCGATCT 1273
Qy 1021 TCTTTACCCACCGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTGTGGCTCTG 1080
Db 1274 TCTATATGCTATGGAAGAGTGTGAAGCTTTCAGTTCCTCAAAATTTATATGTGTGGCTCTG 1333
Qy 1081 CATGTTCTATGCTTTTCCACCTTGTGTTAAATATCTCGCAGAGCTTCTCGATTTGG 1140
Db 1334 CATGTTCTATGCTTTTCCACCTTGTGTTAAACATATTTGCGCAGAGCTTCTGCTTCGG 1393
Qy 1141 TGATCGTGAATTTCTACAAAGGATGTTGGTGAATGCCAAATCTCGAAGATTTATTTGGAGAT 1200
Db 1394 GGATCGTGAATTTCTACAAAGATGTTGGTGAATGCCAAAGTGTGGAGATTTACTGGAGAT 1453
Qy 1201 GTGGAATATGCTGTTCACAAATGATGATCGGCACCTATATTTTCCATGTTTAAAGCA 1260
Db 1454 GTGGAATATGCTGTTCACAAATGATGATGTTTCCATGATATATCTTCCGCTGCTGGCAG 1513
Qy 1261 CGGTCTACCAAGGCTGCTCTTTAAATGCTCTGCTTCTGCTTTCTGTTTATTCCTATGA 1320
Db 1514 CAAGATACCAAGAGACCTCGGCATATATCAATGCTTCTGCTCTGCTGCTGCTTTCATGA 1573
Qy 1321 GCTGTGATGCTGCTTCTGCTGATATTTCAAGTGTGGCTTCTGCTGCTGCTGCTGCTGCT 1380
Db 1574 GCTATGATGCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCT 1633
Qy 1381 TCAGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1634 TCAGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1690
Qy 1441 TGGAAATATGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1691 GGGNACATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1750
Qy 1501 ATACTACCATGCTGATGAATAGGAAAGGCAAACTTGACTGACGTACGG 1551
Db 1751 TTATTACCACGCTGATGAACCGAAAGGATGATGCTGATGCTGATGCTGATGCTGATGCT 1801

RESULT 8

AY054480
LOCUS 2005 bp mRNA linear PLN 05-SEP-2001
DEFINITION Arabidopsis thaliana diacylglycerol O-acyltransferase (At2g19450; F3p11.5) mRNA, complete cds.
ACCESSION AY054480
VERSION AY054480.1 GI:15450799
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 2005)
Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,

Kawai, J., Kim, C., Lin, J., Liu, J., Liu, X., Narusaka, M., Pham, P.K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
Submitted (28-AUG-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
e-mail for correspondence: arab@sequence.stanford.edu

TITLE
JOURNAL

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN Arabidopsis Full-length cDNA"): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, J., Liu, X., Pham, P.K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source

1. 2005
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="II"
/clone="RAFL07-14-014"
/note="This clone is in pBluescript"

gene

1. 2005
/gene="At2g19450; F3p11.5"

CDS

162..1724
/gene="At2g19450; F3p11.5"
/codon_start=1
/product="diacylglycerol O-acyltransferase"
/protein_id="AAK96671.1"
/db_xref="GI:15450800"

/translation="MAILDSAGVTTTENGGEFVDLRLRRKSRSDSSNGLLSGS
DNPSDDVGPADVRDRIDSIVYNDQAGTANLAGNNGDNGGGRGGGRGNAD
ATITYRPSPAHRRARSPLESDAIFRQSHAGLNLVLIANVLIENLMLKYGW
LIRDTWFSSRLDRNPCLCSLIFPLAFTVEKLVIQYSEVPVFLHIITM
TEVLYPVYTLRCDSPAGTGLMLTLCIVLKLVSATYSDIRLSMAADKANPEV
SYVSLKSLAYENVAPTCYQPSYRACIKGKWAQAFKLVITFGMCFITQYIN
PIVNSKHPLKGLDLLVAIERVLKSLVNLVWLCMFVCFELNLIALLLCGDFREF
YKDWNAKSVGDYHRMNMVPHKVRHIYFPCLRKIPKTLAIITAFIAPVAFVHELC
IAPVCLRLFLKWAFLGINFQVPLVITNYIQERFEGSTVGNMIFWIFCFIGQPMCVLLY
YHDLNMRKSGMS"

BASE COUNT 499 a 413 c 447 g 646 t
ORIGIN

Query Match 37.1%; Score 719.8; DB 8; Length 2005;
Best Local Similarity 73.2%; Pred. No. 1e-155;
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;

Qy 241 ACAAGACACTGATTTCTCCGCTCCCTCAAAATTCGCTACCGTCTTCCGCTCCCGCTCACCG 300
Db 446 AGAAGGAGAGGAAGAACCGCGATGCTAGCTTACGTATCGACCGCTCGGTTCAGCTCATCG 505
Qy 301 CAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCCGTCAGAGTACACGGGCGCT 360
Db 506 GAGGCGAGAGAGAGTCCACTTAGCTCCGACCAATCTTCAACAGAGCCATCGCGGATT 565
Qy 361 CTTCACCTTGTATAGTACCTTGTGCTGTAATAGCGGACTCATTCATTGAGAATTT 420
Db 566 ATTCAACCTCTGTGTAGTAGTCTTATTTGCTGTAACAGTAGTACTCATTCGAAAAATCT 625
Qy 421 AATGAAGTATGTTGGTTGATCAAAATCTGGCTTTTGGTTTAAAGTCAATGAGAGA 480
Db 626 TATCAAGTATGTTGGTTGATCACAACGGATTTCTGTTTCTGTTTCAAGATCCCTGCCGAGA 685

```
QY 481 CTGGCCCTTTTCATGTTGTTCTTTCTCTGTTGGTATTATTCCTTTCCGTCGCCCTTTATAGT 540
Db 686 TTGGCCGCTTTTCATGTTGTTATATCCCTTTTCGATCTTTTCCCTTTGGCTGCTTTACGGT 745
QY 541 GGAGAAAGTTGGCAACAGGAAGTGTATACCCGAACAGTGTGTGTGTGTACTTTCATATAAT 600
Db 746 TGAGAAATGGTACTTCAGAAATACATATCAGAACCTGTGTGATCTTCTTCATATAT 805
QY 601 CATTACCTCAACTTCCTTTTCTATCCAGTTTATAGTTTATTTCTCAGGTGTGATTTCTGCTTT 660
Db 806 TATCACCATCAGACAGAGTTTGTATCCAGTTTACGTCAACCTTAAAGTGTGATTTCTGCTTT 865
QY 661 TGTATCAGGTGTACACCTTAACTGTTTCTGTTGTTGTATGTTTAAATTTGGTGTCTTA 720
Db 866 TTTATCAGGTGTACCTTTGATGCTCTCACTTGTGTTGTTGTTGTTGTTGTTTCTTA 925
QY 721 TGCACATACAAACTATGATGAGACACTTACCAATTTAGTTGAAAGGGAGAACCACT 780
Db 926 TGCTCATCTAGTATGACATAGATCCCTAGCCATGCGCTGATAGG----- 975
QY 781 GCTCGATCTCTGAACATGGACTATCCTTTACAACGTAAGCTTCAAGAGCTTGGCATATTT 840
Db 976 -----CCAATCCTGAAGTCTTCTACTAGTTAGCTTGAAGAGCTTGGCATATTT 1024
QY 841 CTTGTTGCTTCCCTACATTTATTTACCAGCCAGCTATCCTCGCACACCTTATATTCGAA 900
Db 1025 CATGGTGGCTCCACATGTTGTTATCAGCAAGTTATCCAGCTTCTGCAATGATACGGAA 1084
QY 901 GGGTGGTGTGTTTCGGCAACTTGTCAAGCTGATAATTTTACAGGAGTTATGGGATTTAT 960
Db 1085 GGGTGGTGGCTGCTCAATTTGCNAACCTGGTCAATTTACCGGATTCATGGGATTTAT 1144
QY 961 AATAGACCAATATTAATCCCAATGATACAAAATTCACAGCATCTCTCTCAAGGGAACCT 1020
Db 1145 AATAGAACAAATATATAATCTTATTTGTCAGGAATCAAAAGCATCTTGAAGGCGATCT 1204
QY 1021 TCTTTTACGCCACGAGAGTCTTGAAGCTTCTGTTTCCAAATTTATATGTTGGGCTCTG 1080
Db 1205 TCTATATGCTATTGAAGAGTGTGGAAGCTTTCAGTTTCCAAATTTATATGTTGGGCTCTG 1264
QY 1081 CATGTTCTATTGCTTTTCCACCTTTGGTTAAATATCTCTGGCAGAGCTTCTTCGATTTGG 1140
Db 1265 CATGTTCTACTGCTTCTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTGCTCTCG 1324
QY 1141 TGATCGTAATTTACAAGATTTGGGAATGCCAAACTGTCGAAGATTTATTTGGAGAT 1200
Db 1325 GCATCGTGAATTTACAAGATTTGGTGAATGCAAAAGCTGCGGAGATTTACGGAAT 1384
QY 1201 GTGGAATATGCTGTTTCACAAATGGATGATCCGCCACCTATATTTTCCATGTTAAGGA 1260
Db 1385 GTGGAATATGCTGTTTCATAAATGGATGTTTCGACATATATATCTCCCGTGTGGCGAG 1444
QY 1261 CGGTCTACCAAGGCTGCTCTCTTTTAAATGCTTCTCTGTTTCTGCTTTATTTCCATGA 1320
Db 1445 CAAGATACCAAGACACTCGCCATTTATCTGTTTCTCTAGTCTCTGCACTCTTTCATGA 1504
QY 1321 CGTGTCATGCTGTTTCTTGGCCACATTTCAAGTTGTGGGCTTTCGGTGGAAATTTATGTT 1380
Db 1505 GCTATGCACTCGCAGTTCCTTGTCTCTCTCTCAAGCTATGGGCTTCTTCTGGGATTTATGTT 1564
QY 1381 TCAGGTTCTCTTGGCTTCTGATCACTAATATCTGCAAAATTAATTTCAAGAACTCAATGTT 1440
Db 1565 TCAGGTTCTCTTGGCTTCTTATCAACAACCTATCTACAGGAAGGTTTGG---CTCAAGGTT 1621
QY 1441 TGGAAATATGATTTTGTGTTTATATTCAGTATCTTGTGTCACACCTATGTTGTGTAAGTCT 1500
Db 1622 GGGGAACATGATCTTCTGTTTCTATCTTCTGCAATTTTCGGNCAACCGTGTGTTGTTCT 1681
QY 1501 ATACTACCATGACTTGTGATGAATAGGAAGCAAACTTGTACTGAAGTACG 1551
Db 1682 TTATTACCAGCACTGATGAACCGGAAAGGATCGATGTCATGAACAACTG 1732
```

```
RESULT 9
AF051849 1845 bp mRNA linear PLN 24-JAN-2000
LOCUS Arabidopsis thaliana diacylglycerol acyltransferase (DAGAT) mRNA,
DEFINITION complete cds.
ACCESSION AF051849 AF061563
VERSION AF051849.1 GI:6625552
KEYWORDS Arabidopsis thaliana.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Bouvier-Nave,P., Benveniste,P., Oelkers,P., Sturley,S.L. and
Schaller,H.
TITLE Expression in yeast and tobacco of plant cDNAs encoding acyl
CoA:diacylglycerol acyltransferase
JOURNAL Eur. J. Biochem. 267 (1), 85-96 (2000)
MEDLINE 20069349
PUBMED 10601854
REFERENCE 2 (bases 1 to 1845)
AUTHORS Benveniste,P.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-1998) Institut de Biologie Moleculaire des
Plantes, Centre National de la Recherche Scientifique, 28 rue
Goethe, Strasbourg 67083, France
COMMENT On Jan 24, 2000 this sequence version replaced gi:6650202.
FEATURES
Location/Qualifiers
source
1..1845
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/tissue_type="hypocotyl"
/dev_stage="3 day-old seedling"
/clone_lib="ABRC library, catalog number CD4-15"
/notes="screened with EST, GenBank Accession Number
AA042298"
1..1845
/gene="DAGAT"
88..1650
/gene="DAGAT"
/function="catalyzes the acylation of diacylglycerol by
oleoyl-CoA to give triacylglycerol"
/notes="acyl-CoA cholesterol acyltransferase homolog"
/codon_start=1
/product="diacylglycerol acyltransferase"
/protein_id="A19262.1"
/db_xref="GI:6625553"
/translation="MAILDSAGVTTVTENGGEFVDLRLRRKSRSDSSNGLLSGS
DNNSPDDVGAPADYDRIDSVYNDDAQTNLANGDNGGNGGNGGGRGGGNAD
ATFTYRSPVARRRARESPSSDAIFKQSHAGLNVVLLVAVNSRLIENLMKYGM
LIRTFWFSSRLRDWPLFMCCISLIPLAATVEKLVOKYISEPVVIFURHITM
TEVLYPVVTLRCDSAFLSGVTMLLTCTVWLKLVYATSVYDILANADKANPEV
SYVSKSLAEVFWAPTLCYQSPRSACIRKQWVARQFAKLVITGTGMLTIEGTIN
PIVRNSKPLKGLLIATLRLVLPNDLVNLYWCMFYCFHMLNLALIEALGQREF
YADWNNAKSGDYRWNNPVMHVMRYIYFPCLRSKPKTLIIATIAVLSVAPHEL
IADPCRLFKLWAFGLGIMFQVPLVITNVLQERFEGSTVGNMFIETFCIFGQPMCVLLY
YHDLNMRKGSMS"
BASE COUNT 453 a 378 c 421 g 593 t
ORIGIN
Query Match 36.8%; Score 715; DB 8; Length 1845;
Best Local Similarity 73.0%; Pred. No. 1.3e-154;
Matches 957; Conservative 0; Mismatches 330; Indels 24; Gaps 2;
QY 241 ACAAGACTGATTTCTCGTCTCTCAAAATCGCTACCGTCTCTCGTCCCCCTCACCG 300
Db 372 AGAAGGAGGAGGAAACCGCGATGCTAGCTTACGTATCGACCGTCTCGTCTCATCG 431
QY 301 CAAAGTGAAGGAAAGTCCGCTCAGCTCCGACACTATTTTCCGTCAGAGTCACCGGCGCT 360
Db 432 GAGGCGAGAGAGATCCACTTAGCTCCGACGCAATCTTCAAAACAGAGCATGCCGATT 491
```

QY 361 CTTCAACCTTTGATATAGTAGTCTTTGTTGCTGTGAATAGCCGACTCATCATTTGAGAATTT 420
Db 492 ATTCAACCTCTGTGTAGTAGTCTTTATGCTGTAAACAGTAGACTCATCATGAAATCT 551
QY 421 AATGAAGTATGTTGGTTCATCAAAATCGGCTTTTGGTTTGTAGTCAAAAGTCAATGAGAGA 480
Db 552 TATGAAGTATGTTGGTTCATCAACAGGATTTCTGGTTTGTAGTCAAAAGTCAATGAGAGA 611
QY 481 CTGGCCCTTTTCATGTGTGTTCTTCTGTGTGATTTCTTCTGCTGCTTTATAGT 540
Db 612 TTGGCCGCTTTTCATGTGTGATATACCTTTTCGATCTTTCTTCTGCTGCTTTACCGT 671
QY 541 GGAGAAGTTGGCACACGAGTAGTATACCCGAACAGTGTGTTGTGACTTCATATAAT 600
Db 672 TGAGAAATTTGGTACTTCAGAAATACATATCAGAACCCTGTGTGATCTTTCTTCATATAT 731
QY 601 CATTAACCTCAACTTCGCTTTTCTATCCAGTATTTAGTATTTCTCAGGTGTGATCTGCTTT 660
Db 732 TATCACCATGACAGAGGTTTGTATCCAGTTTACGTACCTTACCCCTAAGGTGTGATCTGCTTT 791
QY 661 TGATCAGGTGTGACGTTAAGTCTTTCTTCTGCTGTGTGATGTTAAATTTGGTCTCTTA 720
Db 792 TTTATCAGGTGTGACTTTGATGCTCTCTCCTCCTGATGCAATGTGTGGCTTAAAGTTGGTTCTTA 851
QY 721 TGCACATACAAATATGATATGATGAGCACTTACCAAATTTAGTTGAAAGGAGAGCACT 780
Db 852 TGCTCATACTAGTATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 901
QY 781 GCTCATACTCTGAACATGGAGTATCTTACACCTTACAGCTTACAGCTTACAGCTTACAGCT 840
Db 902 -----CCAACTCTGAAGTCTCTACTACTGCTTACCTTACCTTACCTTACCTTACCTT 950
QY 841 CTGTTGCTGCTTACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 951 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010
QY 901 GGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
Db 1011 GGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1070
QY 961 AATGACCAAT 1020
Db 1071 AATGACCAAT 1130
QY 1021 TCTTTACGCCACCGAGAGTCTGAAAGTCTTCTGTTCCAAATTTATATGTTGCTGCTG 1080
Db 1131 TCTATATGCTATTTGAAGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAG 1190
QY 1081 CATGTTCTATTTGCTTTTCCACCTTTGGTTAAATATCTCGGAGAGCTTCTTCGATTTGG 1140
Db 1191 CATGTTCTATGCTTTTCCACCTTTGGTTAAATATCTCGGAGAGCTTCTTCGATTTGG 1250
QY 1141 TGATCGTGAATTTCTAAGAGTGTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1200
Db 1251 GATCGTGAATTTCTAAGAGTGTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1310
QY 1201 GTGGAATATGCTGTTTCAAAATGATGATCGGCCACCTATATTTTCCATGTTTAAAGCA 1260
Db 1311 GTGGAATATGCTGTTTCAAAATGATGATCGGCCACCTATATTTTCCATGTTTAAAGCA 1370
QY 1261 CGGTCTACCAAGGCTGCTGCTTTTAAATGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTG 1320
Db 1371 CAAGATACCAAAACACCTCGGCATATATATGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 1430
QY 1321 GCTGTGATGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 1431 ACTATGCAATCGAGTCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCT 1490
QY 1381 TCAGGTTCTTTGCTTGTGATCACTAATATATCTGCAAAATAAATTCAGAACTCAATGGT 1440
Db 1491 TCAGGTTCTTTGCTTGTGATCACTAATATATCTGCAAACTATCTACAGGAAGGTTGG - -CTCAACGGT 1547

QY 1441 TGGAAATATGATTTTGTGTTTCTATATTCATGATCCTTGGTCAACCTATCTGTGTACTGCT 1500
Db 1548 GGGGAACATGATCTTCTGCTGTTCTCTGATCTTTTCGACAAACGATGTGTGTCTTCT 1607
QY 1501 ATATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551
Db 1608 TTATTACCACGACCTGATGAACCGAAAGGATGATGATGATGATGATGATGATGATGATG 1658
RESULT 10
AF251794 1537 bp mRNA linear PLN 16-APR-2000
LOCUS Brassica napus putative diacylglycerol acyltransferase mRNA,
DEFINITION complete cds.
ACCESSION AF251794
VERSION AF251794.1 GI:7576940
KEYWORDS
SOURCE Brassica napus.
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 1537)
AUTHORS Brown,A.P., Schierer,T.P. and Slabas,A.R.
TITLE Characterization of a putative diacylglycerol acyltransferase cDNA
from Brassica napus embryo
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1537)
AUTHORS Brown,A.P., Schierer,T.P. and Slabas,A.R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2000) Biological Sciences, University of Durham,
South Road, Durham DH1 3LE, UK
FEATURES
1. 1537
source /organism="Brassica napus"
/cultivar="Commanche"
/db_xref="taxon:3708"
/dev_stage="embryo"
22. 1527
/EC_number="2.3.1.20"
/codon_start=1
/product="putative diacylglycerol acyltransferase"
/protein_id="AAF64065.1"
/db_xref="GI:7576941"
/translation="MEILDGGVTMPTEGGADLDTLRHRKPRSDSSNGLLPDSVTYS
DADVRDVSADVEDTCKANLAGENIRESGEAGNVDRVYTRPVSVAHRRVRES
LSSDAIFKQSHAGLNFVVLAVNSRLIENLMKYLRTDFWSSFLRDWPLF
MCCLSLIFPLAFTVEKLVOKISBPVILHIIITMEVLTPVTUUCDSAFUS
GVTLMLTCTIVLWLVSTAHNTYDIRTLANSDDKANPEVSYSLKSLAYFMLAPTLC
YQPSYPRSPCIRKGVARQFAKLIIFTGFMGFIIEQYINPIVRNKRHPLKGLDLYGVE
RVLLKSVNLYVLMCFYCFHLMNLIALLELLCFGDFEYKDMNNAKSVGDYWMNN
PVHKMVRHVVPCLRRNIPKPAIILAFILVSAVHELCIAPVCLPLKLAFLGIMFQ
VPLVFTITLQEREGSMVGNMIFWFTFCIGQPCVCLLYHDLNRRKGRMS"
BASE COUNT 371 a 322 c 355 g 489 t
ORIGIN
Query Match 36.1%; Score 701.2; DB 8; Length 1537;
Best Local Similarity 73.5%; Pred. No. 1.9e-151;
Matches 933; Conservative 0; Mismatches 313; Indels 24; Gaps 2;
QY 264 TCAAAATTCGCTACCGTCTTCCGTCGCCGCTCACCGCAAAGTGAAGAAAGTCCGCTCA 323
Db 272 TAAGTACACGATACGGCGCTCGGTTCAGCTCATCGGAGGTGGGAGAGTCCACTCA 331
QY 324 GCTCCGACATATTTCCGTCAGATGACGCGGCTCTTCAACCTTTGTATAGTAGTCC 383
Db 332 GCTCTGACGCCATCTTCAACAGAGCCCATGCTGGACTATTCAACCTGTGTAGTAGTTC 391
QY 384 TTGTTGCTGTGAATAGCGGACTCATATTGAGAATTTAATGAAGTATGTTGGTTGATCA 443
Db 392 TTGTTGCTGTAAACAGTAGACTCATCATCGAAATCTCATGAAGTACGGTTGGTTGATCA 451
QY 444 AATCTGGCTTTTGGTTTATGTTCAAAAGTCATTGAGAGACTGGCCCTTTTTCATGTTGTC 503

[illegible]

RESULT	13
AX090350	
LOCUS	AX090350
DEFINITION	Sequence 43 from Patent WO0116308.
ACCESSION	AX090350
VERSION	AX090350.1 GI:13444211
KEYWORDS	.
SOURCE	soybean.
ORGANISM	Glycine max Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurousids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	
REFERENCE	1 (bases 1 to 234)
AUTHORS	Lassner,M. and van Eenennaam,A.
TITLE	Plant sterol acyltransferases
JOURNAL	Patent: WO 0116308-A 43 08-MAR-2001;
FEATURES	MONSANTO COMPANY (US)
source	Location/Qualifiers 1. .234 /organism="Glycine max" /db_xref="taxon:3847"
unsure	1. .234 /note="n-unknown"
BASE COUNT	69 a 53 c 40 g 69 t 3 others
ORIGIN	

[illegible][illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 275)
Lassner M. and van Eenennaam, A.
Plant sterol acyltransferases
Patent: WO 0116308-A 45 08-MAR-2001;
MONSANTO COMPANY (US)

Glycine.

```

FEATURES
  source
    1. .275
      location/qualifiers
        /organism="Glycine max"
        /db_xref="taxon:3847"
    1. .275
      /note="n-unknown"
  BASE COUNT
    59 a 56 c 62 g 84 t 14 others
  ORIGIN

```

	Query Match	10.08;	Score 194.6;	DB 6;	Length 275;
	Best Local Similarity	82.2%;	Pred. No. 2.3e-34;		
	Matches 226;	Conservative	0;	Mismatches 48;	Indels 1; Gaps 1;
QY	1165	GTGGAATGCCAAAAC	TCTGCAAGATATT	TGCGAGGATCTGGAATAT	GCCTGTTCACAATG 1224
Db	1	GTGGAATGCCAAAAC	TCTGCAAGATATT	TGCGAGGATCTGGAATAT	GCCTGTTCACAATG 60
QY	1225	GATGATCGGCACCTAT	ATTTTCCATGCTTT	TAAGSCACGGTGCT	TACCAAGCGTCGTGCTCT 1284

Search completed: February 21, 2003, 17:47:13
Job time : 5051 secs